NPS55-90-20

NAVAL POSTGRADUATE SCHOOL
Monterey, California

LATENT FACTOR MODELS AND ANALYSES
FOR OPERATOR RESPONSE TIMES

Donald P. Gaver
I. G. O'Muircheartaigh

September 1990

Approved for public release; distribution is unlimited.

Prepared for:
Naval Postgraduate School,
ey, CA 93955

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. W. West, Jr.
Superintendent

Harrison Shull
Provost

This report was prepared by:

fied

REPORT DOCUMENTATION PAGE

ecurity Classification UNCLASSIFIED	1b Restrictive Markings			
Classification Authority	3 Distribution Availability of Report			
sfication/Downgrading Schedule	Approved for public release; distribution is unlimited			
g Organization Report Number(s) NPS55-90-20	5 Monitoring Organization Report Number(s)			
Performing Organization	6b Office Symbol	7a Name of Monitoring Organization		
tgraduate School	(If Applicable) OR			
(city, state, and ZIP code)	7b Address (city, state, and ZIP code)			
, CA 93943-5000				
Funding/Sponsoring Organization	8b Office Symbol	9 Procurement Instrument Identification Number		
Postgraduate School	(If Applicable)	0&MN, Direct Funding		
(city, state, and ZIP code) Monterey, California	10 Source of Funding Numbers			
ey, CA 93943	Program Element Number	Project No	Task No	Work Unit Accession No

clude Security Classification) Latent Factor Models and Analyses for operator Response Times				
Author(s) Gaver, D. P. and O'Muircheartaigh, I. G.				
of Report	13b Time Covered	14 Date of Report (year, month,day)		15 Page Count
al	From To	1990, September		

nentary Notation The views expressed in this paper are those of the author and do not reflect the official position of the Department of Defense or the U.S. Government.

Codes		18 Subject Terms (continue on reverse if necessary and identify by block number)
up	Subgroup	Latent factor models; extreme-value distribution; Weibull distribution; bootstrap; maximum likelihood estimation; operator response times

ct (continue on reverse if necessary and identify by block number

models are presented for the response times of different operators to different tasks where response is by one or more cues provided by the system. One model for the log-response times is a mixed or latent del with unequal case fixed effects and variances. The other model for the log-response times is a non-log-extreme-value model. Procedures for estimating the parameters by maximum likelihood are l. The models are used to analyze response time data from simulator experiments involving nuclear ant operators performing certain safety-related tasks. The findings of the models are critiqued and ons to risk analysis are sketched.

bution/Availability of Abstract	21 Abstract Security Classification	
lassified/unlimited ☐ same as report ☐ DTIC users	Unclassified	
e of Responsible Individual	22b Telephone (Include Area code)	22c Office Symbol
aver	(408) 646-2605	OR/Gv

1 1473, 84 MAR   83 APR edition may be used until exhausted   security classification of this page

All other editions are obsolete   Unclassified

LATENT FACTOR MODELS AND ANALYSES FOR OPERATOR RESPONSE TIMES

D. P. Gaver
Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943

I. G. O'Muircheartaigh
University College
Galway, Ireland

0. INTRODUCTION, BACKGROUND, AND SUMMARY

There are many situations in which an operator (single individual, or group or crew) is confronted with a somewhat complex task that must be accomplished within prescribed time limits. The task actually often initially requires diagnostic steps followed by action. In some cases the diagnostic steps are stimulated by a cue event, leading to probing actions intended to reveal the correctness of a tentative diagnosis, followed by observation and interpretation of system response, in turn followed by viewpoint revision and further action. While it is intriguing to attempt to model response in such detailed terms, this paper does not embark on that enterprise. Rather, we provide and analyze models for the resulting overall response time of different operators to different tasks where response is initiated by one or more cues provided by the system. Two factor-analytic models are presented along with likelihood estimation procedures. The latter are then employed to analyze data sets from typical exercises conducted at simulators used for training nuclear power plant operators; their identities are kept anonymous. (The findings of the model are critiqued, and applications to risk analysis are sketched.)

1

It is believed that similar models will be useful for summarizing the behavior of operators or crews in other situations, both military and otherwise. For example, application to military tank driver performance is envisioned.

1. A MIXED OR LATENT FACTOR MODEL WITH THE UNEQUAL CASE FIXED EFFECTS AND VARIANCES (LOG N MODEL).

Consider this linear model of mixed type:

$$Y_{ik} = \mu + \nu_k + \omega_i + \epsilon_{ik} \qquad \begin{aligned} i &= 1, 2, \ldots I \\ k &= 1, 2, \ldots K. \end{aligned} \tag{1}$$

where $Y_{ik} = \ln T_{ik}$ with T_{ik} being the time for crew i to respond to situation k; μ and ν_k are fixed constants (effects), and $\omega_i \sim \text{IIDN}(0, \sigma_\omega^2)$, $\epsilon_{ik} \sim \text{IIDN}(0, \sigma_k^2)$, are, respectively, the latent random component that "individualizes" *case* (individual, crew, etc.) i, and the random variation displayed by *any* individual on *situation* (task, problem, etc.) k. It is assumed that each case occurs in conjunction with each situation (e.g. a person confronts a particular problem) just once in the data set to be modelled. In practical circumstances, some such individual interactions may be missing for reasons unrelated to individuals and situations, a problem that is deferred for the present; see Appendix A and B.

As implied, the model described may well be of interest when data pertaining to human performance are to be analyzed, but should also be of use elsewhere. The K tasks or items are allowed to have their own fixed response properties, described by (ν_k, σ_k^2); this pair will be referred to as a *task signature*. The usual mixed ANOVA model formulation assumes $\sigma_k^2 = \sigma^2$, constant for all k (see Scheffé (1959)), as is reasonable when measurement error is represented.

Note that because of the assumption of possibly unequal σ_k's a fixed-ω_i model cannot be usefully estimated by likelihood. Consequently the above random-effect model has been introduced, and fitted to data. As will be apparent, it is possible to estimate the posterior density for ω_i using Bayes' formula in the style of empirical Bayes; the mean of the resulting Normal/Gaussian density $\hat{E}[\omega_i|\text{data}, \log \text{N model}]$, is available as an estimate of the ith crew effect if so desired.

That the above setup is a latent factor model has been remarked to us by Professor T. W. Anderson; see Anderson (1988) Chapter 14 for relevant

3

coverage in the Normal/Gaussian case. Brillinger and Preisler (1983) survey various other latent factor data-analytical studies in non-Gaussian settings; this includes a detailed discussion of a latent factor Poisson model for counting data. Brillinger's paper is interesting in that it suggests examining goodness of model fit by "uniform residuals," a procedure considered in our study as well.

Fitting the LOG N model by likelihood requires iterative calculations; the setup is described in the next section. In case one wishes to "robustify" the formulation, perhaps by introducing more outlier-prone specifications such as the Student t or Tukey density for ω_i then more numerical effort, or approximation is required. Use of the Laplace approximation together with Gauss-Hermite integration may well turn out to be useful; see Gaver and O'Muircheartaigh (1987), and Gaver, Jacobs and O'Muircheartaigh (1990). In a later section a totally non-Normal/Gauss model for operator response times is introduced and fitted.

2. FITTING THE LOG–NORMAL MODEL BY MAXIMUM LIKELIHOOD

In the model (1) the individualizing case effect, ω_i for case i is viewed as a latent or unobserved rv whose effect on the Y_{ik} observable is indirect. What is the probability distribution of $Y_{ik}, k = 1, 2, \ldots K$ in terms of the unknown parameters? Clearly it is multivariate normal since ω_i and ϵ_{ik} occur as a sum; the density for case (crew) i is, by conditional independence, given ω_i ,

$$
\begin{aligned}
f_{\underline{Y}_i}(\underline{y}_i; \mu, \underline{\nu}, \omega_i) &= \prod_{k=1}^{K} \frac{e^{-\frac{1}{2}(y_{ik} - \mu - \nu_k - \omega_i)^2 / \sigma_k^2}}{\sqrt{2\pi}\, \sigma_k} \\
&= \frac{e^{-\frac{1}{2} \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k - \omega_i)^2 / \sigma_k^2}}{(\sqrt{2\pi})^K \prod_{k=1}^{K} \sigma_k}
\end{aligned}
\tag{2}
$$

To obtain the unconditional density of $\underline{y}_i$ remove the condition on ω_i:

$$
f_{\underline{Y}}\left(\underline{y}_i; \mu, \underline{\nu}\right) = \int_{-\infty}^{\infty} f_{\underline{Y}_i}\left(\underline{y}_i; \mu, \underline{\nu}, \omega\right) e^{-\frac{1}{2}\omega^2 / \sigma_\omega^2} \frac{d\omega}{\sqrt{2\pi}\, \sigma_\omega}
\tag{3}
$$

The calculation needed ("completion of the square" in the exponent) can be expeditiously performed as follows. Recognize that the exponent is quadratic in ω; put

$$
-\frac{1}{2}\left(\frac{\overline{\omega}_i - \omega}{\tau}\right)^2 - \frac{1}{2}\, K_i = -\frac{1}{2} \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k - \omega)^2 / \sigma_k^2,
\tag{4}
$$

K_i being independent of ω. To find $\overline{\omega}_i, \tau^2$, and K_i differentiate (2.3) re ω and equate coefficients of ω and 1:

$$
\frac{(\overline{\omega}_i - \omega)}{\tau^2} = \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k - \omega) / \sigma_k^2,
\tag{5}
$$

so, from the ω-term,

$$
1/\tau^2 = \sum_{k=1}^{K} 1/\sigma_k^2
\tag{6}
$$

while from the 1-term

$$\bar{\omega}_i/\tau^2 = \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k)/\sigma_k^2, \tag{7}$$

giving

$$\bar{\omega}_i = \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k) W_k \equiv y_i. - (\mu + \nu.) \tag{8}$$

where $W_k = (1/\sigma_k^2)\,\tau^2 = (1/\sigma_k^2)/\sum_{l=1}^{K}\,1/\sigma_l^2$; $y_i.$ and $\nu.$ are thus W_k–weighted averages. Substitution into (4) gives

$$
\begin{aligned}
K_i &= \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k - \bar{\omega}_i)^2/\sigma_k^2 \\
&= \sum_{k=1}^{K} \left[y_{ik} - \mu - \nu_k - \sum_{k=1}^{K} (y_{ik} - \mu - \nu_k) W_k \right]^2 1/\sigma_k^2 \\
&= \sum_{k=1}^{K} [(y_{ik} - y_i.) - (\nu_k - \nu.)]^2\, 1/\sigma_k^2.
\end{aligned}
\tag{9}
$$

Now from (2)

$$f_{\underline{Y}_i}(\underline{y}_i; \mu, \underline{\nu}, \underline{\sigma}^2, \sigma_\omega^2) = e^{-\frac{1}{2}K_i} \int_\infty^\infty \frac{e^{-\frac{1}{2}(\bar{\omega}_i - \omega)^2/\tau^2}\, e^{-\frac{1}{2}\omega^2/\sigma_\omega^2}}{\left(\sqrt{2\pi}\right)^K \left(\prod_{k=1}^{K} \sigma_k\right) \sqrt{2\pi}\; \sigma_\omega}\; d\omega, \tag{10}$$

a familiar convolution, from which

$$f_{\underline{Y}_i}(\underline{y}_i; \mu, \underline{\nu}, \underline{\sigma}^2, \sigma_\omega^2) = \frac{e^{-\frac{1}{2}K_i}\, e^{-\frac{1}{2}(\bar{\omega}_i)^2/(\tau^2 + \sigma_\omega^2)}}{\left(\sqrt{2\pi}\right)^K \left(\prod_{k=1}^{K} \sigma_k\right) \sqrt{2\pi}\; \sqrt{\tau^2 + \sigma_\omega^2}}\, \tau. \tag{11}$$

Thus the likelihood of $\mu, \underline{\nu}, \underline{\sigma}^2, \sigma_\omega^2$ is proportional to

$$L\left(\mu, \underline{\nu}, \underline{\sigma}^2, \sigma_\omega^2; \underline{y}\right) = \prod_{i=1}^{I} \frac{e^{-\frac{1}{2}K_i}\, e^{-\frac{1}{2}(\bar{\omega}_i)^2/(\tau^2 + \sigma_\omega^2)}}{\left(\prod_{k=1}^{K} \sigma_k\right) \sqrt{\tau^2 + \sigma_\omega^2}}\, \tau \tag{12}$$

6

and hence the log–likelihood, l, may be expressed as

$$2\,l(\mu,\underline{\nu},\underline{\sigma}^2,\sigma_\omega^2;\underline{y}) = I\ \ln\tau^2 -$$

$$I\sum_{k=1}^{K}\ln\sigma_k^2 - I\ \ln(\tau^2+\sigma_\omega^2) - \sum_{i=1}^{I}\ K_i\ -\sum_{i=1}^{I}\ (\bar{\omega}_i)^2/(\tau^2+\sigma_\omega^2) \tag{13}$$

Differentiation gives these estimates:

$$\frac{\partial l}{\partial(\nu_k-\nu.)} = 0 = \sum_{i=1}^{I}\ [(y_{ik}-\bar{y}_i.)-(\nu_k-\nu.)]\left(1/\sigma_k^2\right), \tag{14}$$

so

$$\nu_k\hat{-}\nu. \ \ = \frac{1}{I}\ \sum_{i=1}^{I}\ y_{ik} - \frac{1}{I}\ \sum_{i=1}^{I}\sum_{k=1}^{K}\ y_{ik}W_k \tag{15}$$

$$\equiv (y._k - \bar{y}..).$$

Next,

$$\frac{\partial l}{\partial(\mu+\nu.)} = 0 = \frac{1}{\tau^2+\sigma_\omega^2}\ \sum_{i=1}^{I}\ (y_i. - (\mu+\nu.)) \tag{16}$$

so

$$\mu\hat{+}\nu. = \bar{y}.. \tag{17}$$

These of course closely resemble conventional ANOVA estimates.

When estimating variances it is convenient to reparameterize in terms of precision: $p_k = 1/\sigma_k^2$, $p = 1/\tau^2 = \sum_{k=1}^{K}1/\sigma_k^2 = \sum_{k=1}^{K}p_k$.

Then

$$\frac{\partial l}{\partial p_k} = 0 = -\frac{I}{p} + \frac{I}{p_k} - \frac{-(1/p)^2}{1/p+\sigma_\omega^2} - \Delta_I^2(k) - \frac{(-1)(-(1/p)^2)}{(1/p+\sigma_\omega^2)^2}(\bar{\omega})^2 \tag{18}$$

where

$$\Delta_I^2(k) = \sum_{i=1}^{I}[(y_{ik}-y_i.)-(y._k-\bar{y}..)]^2 \tag{19}$$

and

$$(\bar{\omega})^2 = \sum_{i=1}^{I}(\bar{\omega}_i)^2. \tag{20}$$

Next

$$\frac{\partial l}{\partial\sigma_\omega^2} = 0 = -I\frac{1}{\tau^2+\sigma_\omega^2} + \frac{(\bar{\omega})^2}{(\tau^2+\sigma_\omega^2)^2}, \tag{21}$$

7

yields

$$\frac{1}{p} + \sigma_\omega^2 = \frac{1}{I} \, (\overline{\omega})^2 = \frac{1}{I} \sum_{i=1}^{I} \, (\overline{\omega}_i)^2 \, ; \tag{22}$$

introduction of (21) into (18) simplifies the latter to

$$\frac{1}{p_k} = \frac{1}{I} \, \Delta_I^2 \, (k) + \frac{1}{p}, \qquad k = 1, 2, \ldots, K. \tag{23}$$

The system (15), (16), (22) and (23) must be solved iteratively. Begin by simply fitting as if $\sigma_k^2 = \sigma^2$ to estimate $\mu(1), \nu_k(1), \omega(1)$, and obtain

$$\hat{\sigma}_k^2(1) = \frac{1}{I-1} \sum_{i=1}^{I} \, (y_{ik} - \hat{\mu}(1) - \hat{\nu}_k(1) - \omega_i(1))^2 \tag{24}$$

from which compute $\widehat{W}_k(2) = (1/\hat{\sigma}_\omega^2(1)) / \left(\sum_{l=1}^{K} \, 1/\hat{\sigma}_l^2 \right)$. Next calculate $\nu_k \hat{-} \nu.(2) \equiv y_{\cdot k}(2) - \overline{y}..(2)$ using $\widehat{W}_k(2)$ in (15), and $\mu \hat{+} \nu(2) \equiv \overline{y}..(2)$ from (16). It is now possible to evaluate $\Delta_I^2(k; 2)$ from (19), and $(\overline{\omega})^2(2)$ from (20), and hence $\hat{p}_k(2)$ and $\hat{p}(2)$ from (23), after which $\sigma_\omega^2(2)$ from (22). Now recompute $\widehat{W}_k(3) = (1/\hat{\sigma}_k^2(2)) / \left(\sum_{l=1}^{K} 1/\hat{\sigma}_l^2 \right) = \hat{p}_k(2)/\hat{p}(2)$, and so repeat iteratively until convergence is achieved. A solution procedure based on Newton–Raphson iteration has also been obtained; agreement of the two procedures is generally good.

3. LOG-EXTREME-VALUE MODEL (THE LOG EV MODEL)

An alternative to the previous model that may be attractive is the following setup:

(a) T_{ik} is distributed according to a two–parameter Weibull; then it follows mathematically that

(b) $Y_{ik} = \ln T_{ik}$ has the extreme–value distribution:
$1 - \exp\left[-\theta_i \exp\left((y_{ik} - \eta_k)/\xi_k\right)\right]$, with probability density function

$$f_{Y_{ik}}(y_{ik}; \eta_k, \xi_k; \theta_i) =$$

$$\exp\left[-\theta_i \exp\left((y_{ik} - \eta_k)/\xi_k\right)\right]\theta_i \exp\left((y_{ik} - \eta_k)/\xi_k\right)\frac{1}{\xi_k} \qquad (25)$$

Note the occurrence of parameter θ_i, which is intended to represent crew effect, i.e. θ_i is a way of individualizing crews comparable to the action of ω_i in the previous model. Values of θ_i are viewed as randomly selected latent factors as were the ω_i values. The nature of the θ_i contribution differs from ω_i in this model: whereas in the LOG N model ω_i acted purely additively (on the log scale) to affect the center (mean of logged response times) in a manner common to all tasks, in the LOG EV model it can be seen that logged times are represented as

$$Y_{ik} = \eta_k + \xi_k[(-\ln\theta_i) + \epsilon_{ik}], \qquad (26)$$

ϵ_{ik} having standardized extreme value df. For the present model, (25) or (26),

$$E[Y_{ik}|\theta_i] = \eta_k - 0.5772\xi_k - \xi_k \ln\theta_i \qquad (27)$$

$$Var[Y_{ik}|\theta_i] = \frac{\pi^2}{6}\xi_k^2 \simeq 1.6449\xi_k^2 \qquad (28)$$

which permit initial parameter estimation by moments and facilitates comparison with the results of alternative models. Expression (26) implies that responses to tasks are affected differentially: the greater the natural variation in performing a task by crews (measured by ξ_k for task k) the greater the "average" effect

9

on task duration due to crew effect. This is a specific form of interaction between crew and task effects that may (or may not) be reasonable in particular circumstances.

Conditional on θ_i, the crew i's response, $\underline{Y}_i$, on K different tasks has joint density function

$$f_{\underline{Y}_i}\left(\underline{y}_i;\ \underline{\eta},\underline{\xi},\theta_i\right) = \prod_{k=1}^{K} f_{Y_{ik}}\left(y_{ik};\ \eta_k,\xi_k;\ \theta_i\right), \tag{29}$$

where conditional independence is assumed. In order to obtain the unconditional joint density of response $\underline{Y}_i$ remove the condition on θ_i by integrating out; this step corresponds to (10). Thus

$$f_{\underline{Y}_i}\left(\underline{y}_i;\ \underline{\eta},\underline{\xi}\right) = E_{\theta_i}\left\{e^{-\theta_i}c_i\theta_i^{K}\right\}d_i, \tag{30}$$

where

$$c_i = \sum_{k=1}^{K} \exp\left((y_{ik} - \eta_k)/\xi_k\right) \tag{31}$$

and

$$d_i = \exp\left(\sum_{k=1}^{K}(y_{ik} - \eta_k)/\xi_k\right)\prod_{k=1}^{K}(1/\xi_k) \tag{32}$$

The above model closely resembles one introduced by Crowder (1985) and Crowder and Kimber (1989). However, ours deals with the log time, and hence is a location–scale model that more closely compares to the additive log–normal model, although the η_k is not generally a mean, nor is ξ_k a standard deviation.

4. GAMMA VARIATION FOR θ

A search for mathematical tractability suggests that variation in θ be described by a gamma density:

$$\theta \sim e^{-\theta/\beta}\frac{(\theta/\beta)^{1/\beta - 1}}{\Gamma(1/\beta)}\cdot\frac{1}{\beta} \equiv Gam\left(\frac{1}{\beta},\frac{1}{\beta}\right) \tag{33}$$

10

so $E[\theta] = 1$ and $Var[\theta] = \beta$, and from which the joint density of observations by crew i is

$$f_{\underline{Y_i}}\left(\underline{y}_i, \underline{\eta}, \underline{\xi}\right) = \frac{\Gamma(K + 1/\beta)}{\Gamma(1/\beta)}\left(\frac{1}{1 + \beta c_i}\right)^{K + 1/\beta} \beta^K d_i. \tag{34}$$

The likelihood associated with I independent crews is

$$L\left(\underline{\eta}, \underline{\xi}, \underline{\beta}; y\right) = \left(\frac{\Gamma(K + 1/\beta)}{\Gamma(1/\beta)}\beta^K\right)^I \prod_{i=1}^{I}\left(\frac{1}{1 + \beta c_i}\right)^{K + 1/\beta} d_i \tag{35}$$

or

$$l = \ln L \quad = IK\ln\beta + I\ln\left(\Gamma(K + 1/\beta)/\Gamma(1/\beta)\right)$$

$$- (K + 1/\beta)\sum_{i=1}^{I}\ln(1 + \beta c_i) + \sum_{i=1}^{I}\ln d_i. \tag{36}$$

After arrangement and re-parameterization so that $\phi_k = \ln\xi_k$ the log-likelihood becomes

$$l = I\sum_{k=0}^{K-1}\ln(k + 1/\beta) - (K + 1/\beta)\sum_{i=1}^{I}\ln(1 + \beta c_i) + \sum_{i=1}^{I}\sum_{k=1}^{K}(y_{ik} - \eta_k)\exp(-\phi_k) - I\sum_{k=1}^{K}\phi_k \quad. \tag{37}$$

5. FITTING THE LOG-EXTREME VALUE (LOG-EV) MODEL BY MAXIMUM LIKELIHOOD

To obtain the maximum likelihood estimates of the parameters we iteratively solve the following equations, for which $k = 1, 2, \ldots, K$ throughout:

$$\partial l / \partial \beta = 0,$$

$$\partial l / \partial \eta_k = 0, \tag{38}$$

$$\partial l / \partial \phi_k = 0.$$

One Newton–Raphson iteration only is applied to each equation, after which the entire process is repeated until convergence. Typically, only two or three repetitions are required.

We record the derivatives needed for the above process.

$$
\frac{\partial l}{\partial \beta} = -(K + 1/\beta) \sum_{i=1}^{I} (c_i / (1 + \beta c_i)) + (1/\beta^2) \sum_{i=1}^{I} \ln(1 + \beta c_i)
$$

$$
-I \sum_{k=0}^{K-1} 1/\beta(1 + k\beta) \tag{39}
$$

$$
\frac{\partial l}{\partial \eta_k} = (K\beta + 1) \exp(-\phi_k) \sum_{i=1}^{I} e^{r_{ik}} / (1 + \beta c_i) - I e^{-\phi_k} \tag{40}
$$

where $r_{ik} = (y_{ik} - \eta_k)/\xi_k$, a residual; finally

$$
\frac{\partial l}{\partial \phi_k} = (K\beta + 1) \sum_{i=1}^{I} r_{ik} e^{r_{ik}} / (1 + \beta c_i) - \sum_{i=1}^{I} r_{ik} - I. \tag{41}
$$

12

The second derivatives are

$$
\frac{\partial^2 l}{\partial \beta^2} = IK/\beta^2 + (K + 1/\beta) \sum_{i=1}^{I} c_i^2/(1 + \beta c_i)^2 + (2/\beta^2) \sum_{i=1}^{I} c_i/(1 + \beta c_i)
$$

$$
-(2/\beta^3) \sum_{i=1}^{I} \ln(1 + \beta c_i) - (I/\beta^4) \sum_{k=0}^{K-1} (k + 1/\beta)^{-2}
$$

$$
-(2I/\beta^3) \sum_{k=0}^{K-1} (k + 1/\beta)^{-1},
$$

$$(42)$$

$$
\frac{\partial^2 l}{\partial \eta_k^2} = (1 + K\beta)e^{-2\phi_k} \sum_{i=1}^{I} \left\{ (e^{r_{ik}}/(1 + \beta c_i)) \left[(\beta e^{r_{ik}}/(1 + \beta c_i)) - 1 \right] \right\}, \quad (43)
$$

$$
\frac{\partial^2 l}{\partial \phi_k^2} = (1 + K\beta) \sum_{i=1}^{I} \left\{ \beta \left(r_{ik} e^{r_{ik}}/(1 + \beta c_i) \right)^2 - e^{r_{ik}} r_{ik}(1 + r_{ik})/(1 + \beta c_i) \right\}
$$

$$
+ \sum_{i=1}^{I} r_{ik}.
$$

$$(44)$$

In order to use the inverse of the Fisher information matrix to provide standard errors of the parameter estimates all cross–partial derivatives are required; we omit recording these in the interest of brevity; the expressions may be obtained from the authors. Our numerical experience has been that standard errors obtained from Fisher information tend to be too small, as judged from bootstrapping approaches next to be described.

6. BOOTSTRAPPING

A modern alternative for obtaining standard errors and approximate confidence limits is the *parametric bootstrap* of Efron (1979; esp. Remark K, p.25). This procedure has recently been applied to failure data in the context of the Challenger disaster by Dalal, Fowlkes and Hoadley (1988) and goes as follows:

13

put $\underline{\theta} = (\phi_\omega^2, \mu, \nu_1, \nu_2, ..\nu_K; \sigma_1^2, \ldots, \sigma_K^2)$ in Model LOG N, and $(\beta; \eta_1, \ldots, \eta_K; \phi_1, \ldots, \phi_K)$ in Model LOG EV. Note that $\underline{\theta} = (\theta_1, \ldots, \theta_p)$ here denotes a generic parameter; it bears no direct relation to the ith crew effect in our LOG EV model, (25). Then our procedure is this:

(a) Estimate $\underline{\theta}$ from data; the result is $\widehat{\underline{\theta}}(0)$, the point estimate of the parameters.

(b) Provisionally adopt $\widehat{\underline{\theta}}(0)$ as the true value in the parametric model, in the present case (1) or (25).

(c) Simulate B independent data sets (bootstrap samples) from the model evaluated at $\widehat{\underline{\theta}}(0)$: $\{Y_{ik}(b), i = 1, \ldots, I, k = 1, 2, \ldots, K; b = 1, 2, \ldots, B\}$.

(d) Compute estimates of $\underline{\theta}$ for each sample, obtaining the bootstrap estimates $\{\widehat{\underline{\theta}}(b), b = 1, 2, \ldots, B\} = \{\widehat{\underline{\theta}}(B)\}$, the bootstrap distribution of $\widehat{\underline{\theta}}$.

(e) Present relevant statistical summaries of marginal and joint distributions of $\{\widehat{\underline{\theta}}(B)\}$: e.g. use as standard error of $\widehat{\underline{\theta}}(0)$ components the corresponding standard deviations of the bootstrap estimate; use as confidence limits upper and lower percent points of the bootstrap sampling distributions, suitably adjusted. We present numerical illustrations in the next section.

(f) The same procedure can evaluate standard errors of, and confidence limits for predictions from data: in the present case prediction of the probability that a response time exceeds any given value is evaluated in terms of the model evaluated repeatedly at bootstrap parameter estimate values; see Dalal *et al* (1988) for an example. See Section 9 for an example in the present context.

Bootstrapping methods suggest themselves for comparing the adequacies of different models for fitting and predicting from specific data sets. Specifically, bootstrapping may assist in choosing between two, or more, candidate models.

In the present setting one may wish to predict the probability of non-success, i.e., of response time exceeding some time window of duration t, $\overline{P}(t; \underline{\theta})$. Models A and B (e.g. our LOG N and LOG EV options) are estimated obtaining $\widehat{\underline{\theta}}_A(0)$ and $\widehat{\underline{\theta}}_B(0)$. Then generate bootstrap samples for A and B using $\widehat{\underline{\theta}}_A(0)$ and $\widehat{\underline{\theta}}_B(0)$ respectively, resampling to estimate the mean–squared error of prediction when Model i is used to predict, given that the data comes from Model j; here $\{(i,j)\} = \{A, A;\ A, B;\ B, B;\ B, A\}$. Prefer the model whose use minimizes the maximum estimated mean–squared error of prediction. An alternative strategy is to prefer prediction from the most conservative model: the one predicting the greatest risk; see Section 9 and Draper, Hodges, *et al* (1987).

Another option for residual examination is to compute estimates of the expected log response times associated with each Task/Crew combination. Since crew effects are random, we estimate them in specific cases by their posterior means.

For the LOG N model, examination of (10) reveals that the posterior density of $\underline{\omega}_i$ is $N\left[(\bar{\omega}_i/\tau^2)/(1/\tau^2 + 1/\sigma_w^2), 1/(1/\tau^2 + 1/\sigma_w^2)\right]$. We substitute in the mle's for the various parameters to estimate in a particular case:

$$\hat{\omega}_i = \left(\bar{\omega}_i/(\tau)^2\right)/\left(1/\tau^2 + 1/\sigma_\omega^2\right)$$

and then from (1), (8), (15), (16)

$$\hat{y}_{ik} = (\mu \hat{+} \nu.) + (\nu_k \hat{-} \nu.) + \hat{\omega}_i = \bar{y}.. + (y._k - \bar{y}..) + \left[(y_{i.} - \bar{y}..)/(\hat{\tau})^2\right]/[1/(\hat{\tau}^2) + 1/\hat{\sigma}_\omega^2]$$

$$(45)$$

where all averages are suitably weighted. Note that the above formula for the mean acts, in effect as if a preliminary hypothesis test for homogeneity of crews is being applied: if σ_ω^2 is very small, giving evidence that all crews are the same, then the estimate $\hat{y}_{ik} \simeq y._k$, the (weighted) task mean for each crew. On the other hand if $\hat{\sigma}_\omega^2$ is very large then $\hat{y}_{ik} \simeq y._k + (y_{i.} - \bar{y}..)$, the task mean modified by the estimated effect for crew i. The effectiveness of such a smooth transition when pooling data was noted by Mosteller (1947).

15

For the LOG EV model take the expectation of (27) with respect to the (estimated) posterior density of θ_1, which is Gamma($c_i + 1/\beta, K + 1/\beta$) from (33) and (34). Using the first two terms of the asymptotic expansion we find

$$
\begin{aligned}
\hat{y}_{ik} &= \hat{E}\left[y_{ik}\right] \\
&= \hat{\eta}_k - 0.5772\hat{\xi}_k - \hat{\xi}_k \left[\ln\left(K + 1/\hat{\beta}\right) - \ln\left(\hat{c}_i + 1/\hat{\beta}\right) - 0.5/\left(K + 1/\hat{\beta}\right)\right]
\end{aligned}
\tag{46}
$$

16

7. EXAMPLE DATA ANALYSES

The previous models have been used to analyze response time data from simulator experiments involving operators performing certain safety-related tasks. In Tables 1 and 2 appear actual data from two such: System L and System D. It is noted that certain task-crew combinations are missing, some because of simulator failure. Two procedures were adopted for dealing with these cases: (1) values were imputed by an EM-like process; see Little and Rubin (1987); alternatively, (2) a likelihood approach was taken that simply omits such values from the analysis by setting to unity the likelihood contribution associated with a cell having a missing response. Both approaches can be useful; the former leans more heavily on model correctness. The analyses reported here emphasize the use of a simple imputation procedure; the incomplete data results are also reported in the summary tables.

In addition to "true" missing values there are observations, here marked with asterisks, that were judged to result from operators following non-standard response strategies. It was judged to be useful to analyze the data both with, and without, including such values; when omitted, those nonentries in the data table were treated as missing values, entries imputed as above or treated as missing values, and analyses made using LOG N and LOG EV models.

Results of fitting, along with bootstrapped standard errors, appear in Tables 3 and 4 for System L, and Tables 5 and 6 for System D. Both tables exhibit main effects ($\mu + \nu_k, \eta_k$) and scale parameters (σ_k and ξ_k) for LOG N/EV models computed under A: missing values, and non-standard strategy values, both treated as literally missing, using methods of the Appendix, and alternatively with entries imputed, and B: only the "true" missing values treated as missing. The imputation process used was iteration based on standard two-way ANOVA with fixed task and crew effects. This is a convenient crude approximation to a proper EM algorithmic approach, Little and Rubin (1987). In addition, random crew effect variance parameters, σ_ω^2 and β respectively for the two models, were

17

estimated. The resampling-refitting parametric bootstrap supplied the standard errors; see Efron (1979) and Dalal, Fowlkes, and Hoadley (1988).

It is noted that for System L the fixed effects $(\mu \hat{+} \nu_k, \hat{\eta}_k)$ under A and B agree closely, with the exception of Tasks 2 and 9. Data for Task 9, in Table 1, exhibits 8 out of 18 missing values, and a further 4 non-standard strategy values, all of which are far in excess of other times for that task. Data for Task 2 also exhibit substantially many missing values **and** non-standard times, the latter having resulted from operators following non-standard procedures and hence yielding times more lengthy than the other, acceptable, values. The noticeable differences are, however, still within 2 bootstrap standard errors. The corresponding scale effects (e.g., log task time standard deviations) for Tasks 2 and 9 behave in corresponding fashion, increasing by factors of 2 to 3 if the non-standard times are included.

Similar behavior occurs for System D, although here the exceptions occur for Tasks 4 and 8. There are fewer missing and non-standard times reported for System D than for System L. Relatively large changes occur in the standard errors, as well as in main effect and scale parameter estimates, when several missing or non-standard times are encountered, and these are treated differently in the analysis.

In order to check for the effect of imputation, and also for that of apparent correlations between certain task times the analyses were re-run for System L omitting Tasks 2, 4, and 9. The results appear in Tables 7 and 8. Although specific numerical values are changed, the general pattern remains quite similar: the new numbers are quite often well within a standard error of the estimates that utilize data from all tasks.

18

8. MODEL CRITICISM VIA RESIDUAL EXAMINATION

In order to examine the overall fit of the models to data it is useful to conduct some form of residual analysis. We have chosen first to judge the overall degree of fit by computing and summarizing *uniform residuals* as described by Brillinger and Preisler (1983). In general for this procedure one estimates model parameters $\underline{\theta}$, from data and then examines the estimated probability integral transformation of the data, utilizing the fitted model: if the model is correct then the latter should closely resemble the uniform distribution. In Figures 1 through 8 we display plots and summary statistics for such estimated probability integral transforms of the present data set. We also exhibit the result of bootstrapping once: each model was allowed to create one set of bootstrap sample data utilizing the fitted parameter values; these values were then treated like raw data, and were then probability integral transformed and the results plotted and summarized.

The results have different implications for the appropriateness of the models for the two data sets. The left uniform plot of Figure 1 shows decided non-uniformity of residuals when the raw data is fitted by LOG N using the methods of the Appendix; however, if a bootstrap sample is generated using the fitted model parameters the results are far more uniform. This strongly suggests that the basic model is inappropriate. A similar implication is obtained by examining Figure 3, the residuals of which are associated with LOG EV. Figures 2 and 4 describe the residuals when imputation of missing and/or non-standard values is conducted. Notice that uniformity of residuals of the fit of the raw (plus imputed) data is greatly enhanced. This is not surprising since imputation is based on presumption of model correctness, and the missing and non-standard values are imputed using the presumed model. The same general behavior is observed when data from System D are fitted by the two models in various ways. Here, however, the departure of the residuals from uniformity, as shown on the left-most residual plots of Figures 5 and 7, seems less pronounced than is the

case for System L. Again, imputation improves the uniformity of the residuals and the apparent fit. We conclude that the log-additive models are more likely to be trustworthy for system D than for System L.

9. MODIFICATION OF THE MODELS FOR INITIAL DELAY

In the nuclear plant simulator exercises, and doubtless for other applications as well, certain task response times may begin after a cue different from, and later than, the actual initiating event. That is, the operational cue that triggers response may occur some time after a possible original cueing event (e.g., the first evidence of nuclear plant abnormality). Unfortunately, the time of the *operational* cue is not usually recorded in simulator practice, and so response time data, which may use initiating event time or some other well-specified event time for reference, tends to exhibit an initial delay. Such delays may be inferred from plant and initiating event information, or by examining the actual response time data. Note that ignoring such delays if they are appreciable, i.e., fitting a 2-parameter LOG N or LOG-EV model when a 3-parameter specification is more appropriate can importantly change the estimated parameter values, particularly the LOG N σ_k or LOG EV ξ_k, the measures of within-task variability. Specifically, if the delay is ignored and our current models employed uncritically when a delay $\gamma_k > 0$ is required then the estimated values, $\hat{\sigma}_k$ and $\hat{\xi}_k$, will be *biased downwards* (under-estimated), sometimes quite significantly.

To illustrate the effect of a rough accounting for delay examine the Task 10 data for system L in Table 1. The minimum value is 1402 and the maximum is 3450, so there is an appreciable delay associated with beginning the task as compared to the variability of the response times. If the data is taken at face value and our LOG N and LOG EV models fitted, then Table 3 exhibits $\xi_{10} = 0.28$ considerably smaller than that for other tasks. If, however, a rough adjustment is made for delay by subtracting 1000 from each Task 10 response time data value then the simple standard deviation estimate for logged (times-1000) for Task 10 becomes 0.59, far more similar to other task response time standard deviations.

At present the LOG N and LOG EV programs fit 2-parameter models, so accounting for delay, γ_k, must be done off-line. A formal approach to the es-

21

timation problem for the Weibull model is given by Smith and Naylor (1987); that paper also references other relevant articles.

10. RISK CALCULATIONS

An important application of the LOG N and LOG EV models is to **risk analysis**: it is desired to estimate the probability that a task's response time occurs within a particular time window, for in that case (some aspect of) the threat has been averted. We will refer to the probability that response time exceeds the time window as the human interaction risk associated with the task.

It is easy to make point estimates of the required probabilities using both existing models: one simply replaces the model parameters associated with the task of interest by their maximum likelihood estimates. A natural way in which to handle crew effect is simply to remove the crew condition ω or θ by integrating it out with respect to the appropriate Normal/Gauss or Gamma estimated prior. In order to assess the effect of a particular crew on the risk it is necessary to calculate the posterior density for that crew and then integrate out on ω or θ with respect to that posterior. It may well be of interest to compare the estimated risk as it depends upon which crew is in place when an initiating event occurs in order to assess the effect of individual crews directly on risk. This is not done here.

11. RISK UNCERTAINTY

In order to assess the the uncertainty inherent in the risk estimation one may once again bootstrap. The procedure is this:

(a) Estimate θ from data to obtain $\hat{\theta}_0$:

$$\text{LOG N}: \ \hat{\theta}_0 = (\hat{\sigma}_\omega^2, \hat{\mu}, \hat{\nu}, \hat{\sigma}^2)$$

$$\text{LOG EV}: \ \hat{\theta}_0 = (\hat{\beta}, \hat{\eta}, \hat{\xi}^2)$$

(b) Provisionally adopt $\hat{\theta}_0$ as the true value in the parametric model.

(c) Simulate B independent data sets from the model with parameter $\hat{\theta}_0$: $\{y_{ik}(b), \text{ with } b = 1, 2, \ldots, B\}$.

(d) Estimate θ from each sample of (c). Obtain $\{\hat{\theta}(b), \ b = 1, 2, \ldots, B\}$;

(e) Compute $P\{T_k > w_k | \theta(b)\} = P\{\ln \ T_k > \ln \ w_k | \theta(b)\}$;

$= P\left\{Y_k > w'_k | \theta(b)\right\} \equiv r_k(b)$, where

$$w'_k = \ln w_k.$$

The risk associated with Task k estimated from bootstrap sample $b(b = 1, 2, \ldots, B)$ for each model. that is, calculate the probability that the time window w_k is exceeded using the bth set of parameters estimated. For our two models and $b = 1, 2, \ldots, B)$, and also $b = 0$, the original estimate, thus becomes, respectively, for the two models,

$$R_k(b; \text{LOG N}) = 1 - \Phi\left(\frac{w'_k - (\hat{\mu}(b) + \hat{\nu}_k(b))}{\sqrt{\hat{\sigma}_\omega^2(b) + \hat{\sigma}_k^2(b)}}\right), \tag{47}$$

$$R_k(b; \text{LOG EV}) = \left(1 + \hat{\beta}(b) \exp\left(w'_k - \hat{\eta}_k(b)\right) / \hat{\xi}(b)\right)^{-1/\hat{\beta}(b)} \tag{48}$$

24

The following table exhibits risk calculation for prescribed windows; the lengths chosen are for illustration only. Calculations have been made both *without* non-standard strategy values (A) and *including* non-standard strategy values (B); any missing or left-out values were imputed as before.

Examination of the point estimates in Table shows that there is considerable similarity in the orders of magnitude of the risks given by LOG N and LOG EV. However, LOG EV values are consistently below those for LOG N, as are the corresponding confidence limits. It is, thus, more conservative to adopt LOG N model-based risk estimates than to use those based on LOG EV, or equivalently the Weibull model. The exception, Task 9, is probably traceable to the many missing values exhibited.

(f) Present statistical summaries of marginal and joint distribution of $R_k(b)$: e.g., use as standard error for the original risk estimate, $R_k(0)$, the standard deviation

$$S_{R_k} = \sqrt{\frac{1}{B-1} \sum_{k=1}^{K} \left(R_k(b) - \overline{R}_k \right)^2} \tag{49}$$

Note that the *overall risk* associated with a particular initiating *event* depends upon the risks associated with all *tasks* (human interactions) associated with response to the event. These risks may well be dependent probabilities, at minimum because all tasks presumably confront the same crew. To handle the dependency induced by a common crew one can assume conditional independence, multiply risks conditional on crew to obtain the risks associated with joint events, and then integrate out with respect to the crew's posterior density. The calculation may be performed explicitly for the LOG EV model since a closed-form elementary function expression for the extreme-value survivor function exists; no such simple calculation can be carried out for the LOG N, but numerical procedures are always available.

25

Examination of the results suggests considerable similarity between the risks calculated using the LOG N and LOG EV models. It is, however, noticeable that for all tasks except Tasks 3 and 9, LOG N predicts a slightly greater risk than does LOG EV, and generally with slightly larger standard error. Of course in most cases shown the risks associated with the window values in our example are too high to be realistic; the windows were chosen for illustration only.

The above calculations have been carried out using parameter values obtained under imputation. Thus the apparent similarity of risks across models is probably overstated, and, in view of the suspicion cast on the fits to system L data (see Section 8) we should treat the System L risks with caution.

12. SUMMARY AND CONCLUSIONS

In this article we have suggested and shown how to fit, by maximum likelihood, two models for operator response times. The fits of the models to two sets of (actual) data are displayed and compared. Uncertainties are assessed by parametric bootstrapping. Complications involving missing values and consequent lack of balance are dealt with by direct likelihood computation as well as by a simple form of imputation. Finally, the fitted models are applied to estimate the risk of exceeding (hypothetical) time windows; associated uncertainties, i.e., standard errors and confidence limits, are obtained by bootstrapping.

We view this work as a pilot or feasibility study intended to illustrate and explore possibly useful approaches and methodology to an important area. Outstanding problems remain: the models put forward were chosen for their abilities to account for some aspects of the real situation and for relative tractability, but many other forms could be conjured up, fitted, and applied to infer risk, as defined here. It is somewhat interesting to find that risks estimated from the same data using the different models agree rather closely; it is not unlikely that the agreement will suffer if the windows are increased so as to achieve much smaller, and presumably more realistic risk values.

The bootstrap standard errors and confidence limits warn that although the present models tend to agree in their risk assessments the uncertainty is still rather large, even if wrong-model or *structural uncertainty* is ignored; see Draper, Hodges *et al* (1987) for discussion. In order to reduce the uncertainty of estimation, e.g. to reduce standard error size, it is often proposed to aggregate or pool data, either from similar tasks in the same environment (plant), or for the same task across "similar" environments. Both procedures are worthy of investigation, but will be credible only if suitable adjustments are made to reduce bias. Adjustments can be carried out by using models resembling the types suggested here, possibly enhanced to include regression terms "explaining" responses in terms of measured and qualitative crew and plant characteristics ("performance

27

shaping factors" is the jargon in certain risk assessment circles). To date, formal adjustment attempts by regression have been inconclusive but are a form of insurance that should be included, and validated to the greatest extent possible, if aggregation or pooling is contemplated, particularly across plants. Inter-plant variability may be appreciable because of variations in management philosophy and style.

In general, it seems advisable to utilize bootstrapping as extensively as possible to build an appreciation for the variabilities and uncertainties involved when using models. Bootstrapping that uses direct re-sampling as originally discussed, Efron (1979), seems difficult or impossible for situations such as are described here unless vastly more data becomes available and better, more scientifically based, models and true replications can be employed. Consequently, use of parametric bootstrapping becomes necessary. However, parametric bootstrapping that consumes data from more realistic, and elaborate, models and their fits to simpler structures can be useful and informative. But such exercises cannot directly substitute for data obtained under truly operational circumstances, which even the best *simulator* data can not aspire to be.

13. ACKNOWLEDGEMENTS

APPENDIX A. MAXIMUM LIKELIHOOD ESTIMATION FOR LOG N MODEL, WITH MISSING VALUES

Let $d_{ik} = 1$ if the observation at (i,k) [i.e., for crew i, task k] is available; otherwise $d_{ik} = 0$. Then the component of likelihood associated with (Crew) i is

$$f_{\underline{y_i}}\left(\underline{y_i}, \mu, \underline{\nu}, \omega_i\right) = \prod_{k=1}^{K} \left\{ \frac{e^{-\frac{1}{2}(y_{ik}-\mu-\nu_k-\omega_i)^2/\sigma_k^2}}{\sqrt{2\pi}\sigma_k} \right\}^{d_{ik}}$$

$$= \frac{e^{-\frac{1}{2}\sum_{k=1}^{K} d_{ik}(y_{ik}-\mu-\nu_k-\omega_i)^2/\sigma_k^2}}{\left(\sqrt{2\pi}\right)^{\sum_k d_{ik}} \prod_{k=1}^{K} \sigma_k^{d_{ik}}}$$

Now write

$$\sum_{k=1}^{K} d_{ik}\left(y_{ik} - \mu - \nu_k - \omega\right)^2 \cdot \frac{1}{\sigma_k^2} = \frac{(\bar{\omega}_i - \omega)^2}{\tau_i^2} + K_i.$$

After differentiation re ω,

$$\sum_{k=1}^{K} \frac{d_{ik}}{\sigma_k^2}\left(y_{ik} - \mu - \nu_k - \omega\right) = \frac{(\bar{\omega}_i - \omega)}{\tau_i^2}$$

which implies, identifying terms of order 1 and ω, and writing

$$\begin{cases} \sum_{k=1}^{K} d_{ik}\left(y_{ik} - \mu - \nu_k\right)p_k = \bar{\omega}_i/\tau_i^2 \\ \sum_{k=1}^{K} d_{ik}p_k = 1/\tau_i^2 \end{cases}$$

Hence $\tau_i^2 = 1/\sum_{k=1}^{K} d_{ik}p_k$

$\bar{\omega}_i = \tau_i^2 \sum_{k=1}^{K} d_{ik}\left(y_{ik} - \mu - \nu_k\right)p_k = y_{i.} - \mu - \tau_i^2 \sum_{k=1}^{K} d_{ik}\nu_k p_k$, where $y_{i.} = \tau_i^2 \sum_{k=1}^{K} d_{ik}y_{ik}p_k$

and

$$K_i = \sum_{k=1}^{K} d_{ik}\left(y_{ik} - \mu - \nu_k - \omega_i\right)^2 p_k.$$

Thus re-write the likelihood as

$$\int \underline{y_i}\left(\underline{y_i}; \mu, \underline{\nu}, \underline{\sigma}, \sigma_\omega^2\right) = e^{-\frac{1}{2}K_i} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left\{(\bar{\omega}_i - \omega)^2/\tau_i^2\right\}} e^{-\frac{1}{2}\omega^2/\sigma_\omega^2} \, dw}{\left(\sqrt{2\pi}\right)^{\Sigma d_{ik}} \sigma_k^{d_{ik}} \sqrt{2\pi}\,\tau_i}$$

$$\propto \frac{e^{-\frac{1}{2}K_i} e^{-\frac{1}{2}\bar{\omega}_i^2/\left(\tau_i^2 + \sigma_\omega^2\right)} \sqrt{\tau_i^2}}{\left\{\prod\limits_{k=1}^{K} \sigma_k^{d_{ik}}\right\} \sqrt{\tau_i^2 + \sigma_\omega^2}}$$

Hence the likelihood assumes the form

$$L = \prod_{i=1}^{I} \frac{e^{-\frac{1}{2}K_i} e^{-\frac{1}{2}\bar{\omega}_i^2/\left(\tau_i^2 + \sigma_\omega^2\right)}}{\prod_{k=1}^{K} \sigma_k^{d_{ik}} \sqrt{\tau_i^2 + \sigma_\omega^2}} \sqrt{\tau_i^2}$$

Taking logarithms, we obtain

$$l = 2\ln L = -\sum_{i=1}^{I} K_i - \sum_{i=1}^{I} \frac{\bar{\omega}_i^2}{\tau_i^2 + \sigma_\omega^2} + \sum_{i=1}^{I} \ln \tau_i^2 - \sum_{k=1}^{K} n_k \ln \sigma_k^2 - \sum_{i=1}^{I} \ln\left(\tau_i^2 + \sigma_\omega^2\right),$$

$$= -\sum_{i=1}^{I} K_i - \sum_{i=1}^{I} \frac{\bar{\omega}_i^2}{\tau_i^2 + \sigma_\omega^2} + \sum_{i=1}^{I} \ln \tau_i^2 + \sum_{k=1}^{K} n_k \ln p_k - \sum_{i=1}^{I} \ln\left(\tau_i^2 + \sigma_\omega^2\right)$$

where

$$n_k = \sum_{i=1}^{I} d_{ik}$$

Differentiating first re μ, we obtain

$$\frac{\partial l}{\partial \mu} = -\sum_{i=1}^{I} \left(\frac{\partial K_i}{\partial \mu} + \frac{2\bar{\omega}_i}{\tau_i^2 + \sigma_\omega^2} \frac{\partial \bar{\omega}_i}{\partial \mu}\right)$$

$$\frac{\partial K_i}{\partial \mu} = 0, \frac{\partial \bar{\omega}_i}{\partial \mu} = -1$$

Hence

$$\frac{\partial l}{\partial \mu} = -\sum_{i=1}^{I} \frac{2\bar{\omega}_i\,(-1)}{\tau_i^2 + \sigma_\omega^2} = 0$$

$$\Rightarrow \sum_{i=1}^{I} \frac{\bar{\omega}_i}{\left(\tau_i^2 + \sigma_\omega^2\right)} = 0$$

31

$$\frac{\left(y_{i.} - \mu - \tau_i^2 \sum\limits_{k=1}^{K} d_{ik} \nu_k p_k \right)}{(\tau_i^2 + \sigma_\omega^2)} = 0$$

$$\hat{\mu} \sum_{i=1}^{I} \left(\tau_i^2 + \sigma_\omega^2 \right)^{-1} = \sum_{i=1}^{I} \frac{\left(y_{i.} - \tau_i^2 \sum\limits_k d_{ik} p_k \nu_k \right)}{(\tau_i^2 + \sigma_\omega^2)}$$

$$\Rightarrow \hat{\mu} = \sum_i \frac{\left(y_{i.} - \tau_i^2 \sum\limits_k d_{ik} \nu_k p_k \right)}{(\tau_i^2 + \sigma_\omega^2)} \Bigg/ \sum_i \left(\tau_i^2 + \sigma_\omega^2 \right)^{-1}$$

Next, differentiate re ν_l:

$$\frac{\partial l}{\partial \nu_l} = -\left\{ \sum_i \frac{\partial K_i}{\partial \nu_l} + \sum_i \frac{2\bar{\omega}_i}{\tau_i^2 + \sigma_\omega^2} \cdot \frac{\partial \bar{\omega}_i}{\partial \nu_l} \right\} = 0$$

$$\frac{\partial K_i}{\partial \nu_l} = \frac{\partial}{\partial \nu_l} \left\{ \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right)^2 p_k \right\}$$

$$= 2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k \left(-\delta_{lk} - \frac{\partial \bar{\omega}_i}{\partial \nu_l} \right)$$

$$\frac{\partial \bar{\omega}_i}{\partial \nu_l} = -\tau_i^2 d_{il} p_l$$

Hence

$$\frac{\partial l}{\partial \nu_l} = -\sum_{i=1}^{I} \left\{ 2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k \left(-\delta_{lk} + \tau_i^2 d_{il} p_l \right) + 2 \frac{\bar{\omega}_i}{\tau_i^2 + \sigma_\omega^2} \left[-\tau_i^2 d_{il} p_l \right] \right\} = 0$$

Thus,

$$-\sum_{i=1}^{I} d_{il} \left\{ \tau_i^2 d_{il} p_l \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k - d_{ik} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right) p_l - \frac{\bar{\omega}_i \tau_i^2}{\tau_i^2 + \sigma_\omega^2} d_{il} p_l \right\} = 0$$

Further simplification results in

$$-\sum_i d_{il} \left\{ \tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k - \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right) - \frac{\bar{\omega}_i \tau_i^2}{\tau_i^2 + \sigma_\omega^2} \right\} = 0$$

Hence

$$\nu_l \sum_l d_{il} = \sum_{i=1}^{I} d_{il} \left[-\tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k + \left(y_{il} - \mu - \bar{\omega}_i \right) + \frac{\tau_i^2 \bar{\omega}_i}{(\tau_i^2 + \sigma_\omega^2)} \right]$$

$$\Rightarrow \nu_l = \frac{1}{n_l} \sum_{i=1}^{I} d_{il} \left\{ -\tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k + \left(y_{il} - \mu - \bar{\omega}_i \right) + \frac{\tau_i^2 \bar{\omega}_i}{(\tau_i^2 + \sigma_\omega^2)} \right\}.$$

We now differentiate re p_l, noting

$$\frac{\partial \tau_i^2}{\partial p_l} = -\frac{1}{\left(\sum_k d_{ik} p_k \right)^2} d_{il} = -\tau_i^4 d_{il}$$

$$\frac{\partial \bar{\omega}_i}{\partial p_l} = \tau_i^2 d_{il} \left(y_{il} - \mu - \nu_l \right) + \left\{ \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k \right) p_k \left(-\tau_i^4 d_{il} \right) \right\}$$

$$= \tau_i^2 d_{il} \left\{ y_{il} - \mu - \nu_l - \tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k \right) p_k \right\}$$

$$= \tau_i^2 d_{il} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)$$

$$\frac{\partial K_i}{\partial p_l} = d_{il} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)^2 + 2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k \left\{ -\tau_i^2 d_{il} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right) \right\}$$

$$= d_{il} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) \left\{ \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right) - 2\tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k \right\}$$

Noting also that

$$\tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k - \bar{\omega}_i \right) p_k = \tau_i^2 \sum_k d_{ik} \left(y_{ik} - \mu - \nu_k \right) p_k - \bar{\omega}_i \tau_i^2 \sum_k d_{ik} p_k = \bar{\omega}_i - \bar{\omega}_i = 0$$

it follows that

$$\frac{\partial K_i}{\partial p_l} = d_{il} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)^2.$$

33

$$\frac{\partial l}{\partial p_l} = \sum_{i=1}^{I} d_{il} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)^2 - \frac{2\bar{\omega}_i \tau_i^2}{\left(\tau_i^2 + \sigma_\omega^2\right)} \left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)$$

$$- \frac{(-1)\bar{\omega}_i^2 \left(1 - \tau_i^2 d_{il}\right)}{\left(\tau_i^2 + \sigma_\omega^2\right)^2} + \frac{1}{\tau_i^2}\left(-\tau_i^4 d_{il}\right) + \frac{n_l}{p_l} - \frac{1\left(-\tau_i^4 d_{il}\right)}{\left(\tau_i^2 + \sigma_\omega^2\right)}$$

Hence,

$$\frac{\partial l}{\partial p_l} = 0 \Rightarrow$$

$$\sum_{i=1}^{I} d_{il} \left\{ -\left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)\left[y_{il} - \mu - \nu_l - \bar{\omega}_i + \frac{2\bar{\omega}_i \tau_i^2}{\tau_i^2 + \sigma_\omega^2} \right] - \frac{\bar{\omega}_i^2 \tau_i^4}{\left(\tau_i^2 + \sigma_\omega^2\right)^2} - \tau_i^2 + \frac{\tau_i^4}{\tau_i^2 + \sigma_\omega^2} \right\} + \frac{n_l}{p_l} =$$

$$p_l = n_l / \left\{ \sum_{i=1}^{I} d_{il} \left[\left(y_{il} - \mu - \nu_l - \bar{\omega}_i \right)\left(y_{il} - \mu - \nu_l - \bar{\omega}_i + \frac{2\bar{\omega}_i \tau_i^2}{\tau_i^2 + \sigma_\omega^2} \right) \right] + \frac{\bar{\omega}_i^2 \tau_i^4}{\left(\tau_i^2 + \sigma_\omega^2\right)^2} + \tau_i^2 - \frac{\tau_i^4}{\tau_i^2 + \sigma_\omega^2} \right.$$

Finally, differentiating re σ_ω^2, we find

$$\frac{\partial l}{\partial \sigma_\omega^2} = \sum_{i=1}^{I} \frac{\bar{\omega}_i^2}{\left(\tau_i^2 + \sigma_\omega^2\right)^2} - \sum_{i=1}^{I} \frac{1}{\left(\tau_i^2 + \sigma_\omega^2\right)} = 0$$

which implies

$$\sigma_\omega^2 = \sum_{i=1}^{I} \frac{\bar{\omega}_i^2 - \tau_i^2}{\left(\tau_i^2 + \sigma_\omega^2\right)^2} / \sum_{i=1}^{I} \frac{1}{\left(\tau_i^2 + \sigma_\omega^2\right)}$$

The following iterative procedure was employed:

(a) Use imputed values to obtain $\hat{\mu}(0), \hat{\sigma}_\omega^2(0), \hat{\gamma}_k(0), \hat{p}_k(0)\,[k = 1, ..., K]$

(b) Solve for $\hat{\mu}(1)$ using (A-1) Update $\bar{\omega}_i$

(c) Solve for $\hat{\nu}_l(1)$ using (A-2) Update $\bar{\omega}_i$

(d) Solve for $\hat{p}_l$ using (A-3) Update $\bar{\omega}_i, \tau_i^2$

(e) Solve for σ_ω^2 using (A-4)

Repeat (b) through (e) until convergence.

34

REFERENCES

Anderson, T. W. (1984) *An Introduction to Multivariate Statistical Analysis* (Second Edition). New York: John Wiley.

Brillinger, D. R., Preisler, H. K. (1983) "Maximum likelihood estimation in a latent variable problem." *Studies in Econometrics, Time Series and Multivariate Statistics.* New York: Academic Press.

Crowder, M. (1985) "A distributional model for repeated failure time measurements." *Journal of the Royal Statistical Society, Ser. B*, 47, pp. 447-452.

Crowder, M. (1989). "A multivariate distribution with Weibull connections." *Journal of the Royal Statistical Society, Ser. B*, 51, No. 1, pp. 93-107.

Dalal, S. R., Fowlkes, E. B. and Hoadley, B. (1990) "Risk Analysis of the Space Shuttle: Pre-Challenger prediction of failure." *Journal of the American Statistical Association*, 84, No. 408, pp. 945-967.

Denby, L., and Mallows, C. (1988) "Robust analysis of mixed models." AT&T Statistical Research Reports, No. 65. *AT&T Bell Laboratories*, Murray Hill, NJ.

Draper, D., Hodges, J. S., Leamer, E., Morris, C. N., Rubin, D. B. (1987). *A Research Agenda for Assessment and Propagation of Model Uncertainty.* Rand Note N-2683-RC. The Rand Corporation, Santa Monica, CA.

Efron, B. (1979) "Bootstrap methods: another look at the jackknife." *Annals of Statistics*, 7, pp. 1-26.

Gaver, D. P., and O'Muircheartaigh, I. G. (1987). "Robust empirical Bayes analyses of event rates." *Technometrics*, 29, No. 1, pp. 1-15.

Gaver, D. P., Jacobs, P. A., and O'Muircheartaigh, I. G. (1990) "Regression analysis of hierarchical Poisson-like event rate data: superpopulation model effect on predictions." (to appear in *Communications on Statistics*).

Hougaard, P. (1986) "Survival models for heterogeneous populations derived from stable distributions." *Biometrika*, 73, pp. 387-396.

Hougaard, P. (1986) "A class of multivariate failure time distributions." *Biometrika*, 73, pp. 671-678.

Kimber, A. C., and Crowder, M. J. (1988) *A Repeated Measurements Model with Applications in Psychology*, University of Surrey Technical Report on Statistics, No. 65.

Little, R. J. A., and Rubin, D. B (1987) *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.

Mosteller, F. W. (1947) "On pooling data." *Journal of the American Statistical Association*, 43, No. 242, pp. 231-242.

Scheffé, H. (1959) *The Analysis of Variance*. New York: John Wiley and Sons.

Smith, R. L., and Naylor, J. C. (1987) "Statistics of the three-parameter Weibull distributions." *Annals of Operations Research*, 9, pp. 577-587.

Watson, A. S. and Smith, R. L. (1985) "An examination of statistical theories for fibrous materials in the light of experimental data." *Journal of Materials Science*, 20, pp. 3260-3270.

TABLE 1
SYSTEM L TASK-CREW TIMES

TASK/CREW	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	87	76	92*	64	134*	9	29	14	59	85	20	7	61	37	48	§	11	11
2	88	370*	93*	65	§	49	30	247*	60	86	31	§	75	38	165*	§	12	12
3	164	34	51	162	109	70	§	25	46	48	67	28	151	34	26	§	20	25
4	§	§	453	261	406	§	§	371	91	876	668	320	§	310	678	§	§	§
5	196	57	§	125	218	260	§	92	65	149	51	110	90	80	62	§	84	14
6	20	16	21	19	21	3	8	12	19	22	20	20	9	9	15	§	§	11
7	§	§	§	§	§	160	171	54	87	75	76	65	68	126	87	§	186	121
8	142	730	314	930	509	324	§	1159	493	256	420	360	449	646	500	§	§	§
9	22	§	721*	2149*	§	1036*	62	10	101	§	§	§	208*	19	37	§	§	§
10	§	1410	1676	§	§	2287	3450	2298	1402	2293	1565	2340	1522	§	1665	§	§	1473
11	§	§	377	89	242	392	802	135	718	246	122	978	40	230	60	535	273	69
12	574	§	2146	§	630	749	3144	1767	840	1004	1246	1055	1419	1888	2578	2338	§	1046

* Alternative strategies
§ Missing data (e.g., simulator error)

TABLE 2.
SYSTEM D TASK-CREW TIMES

TASK/CREW	1	2	3	4	5	6	7	8	9	10
1	10	8	5	6	13	14	10	13	15	24
2	404	263	437	115	403	248	346	168	384	193
3	326	457	272	562	409	401	303	320	283	295
4	431	143*	147	210	1329	§	519	290	31*	§
5	1799	1226	1040	1628	855	439	1225	548	912	§
6	304	276	293	209	276	307	271	510	300	295
7	131	18	33	171	121	141	112	28	71	69
8	200	312	41	262*	250	809*	526*	238	211	868*
9	277	475	229	492	416	848	358	279	936	446
10	420	1192	817	962	1456	1487	690	955	1102	1020

* Alternative strategies

§ Missing data (e.g., simulator error)

MODEL: Log EV

TASK	TASK EFFECTS η_k	STD. ER.	TASK "STD. DEVS." ξ_k	STD. ER.	WINDOW	TASK RISKS RISK	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
1	3.16(3.54)	0.30(0.18)	0.67(0.58)	0.14(0.09)	80.	0.061(0.080)	0.019(0.031)	0.104(0.129)	0.036(0.036)
2	3.04(3.79)	0.44(0.11)	0.66(0.39)	0.20(0.07)	600.	0.000(0.045)	0.002(0.018)	0.005(0.000)	0.000(0.000)
3	3.69(4.01)	0.29(0.17)	0.62(0.41)	0.14(0.10)	160.	0.039(0.045)	0.015(0.018)	0.066(0.071)	0.023(0.022)
4	5.24(5.97)	0.57(0.13)	0.69(0.41)	0.27(0.07)	100.	0.701(0.966)	0.079(0.035)	0.831(1.007)	0.600(0.904)
5	4.24(4.62)	0.28(0.17)	0.55(0.49)	0.11(0.11)	200.	0.061(0.089)	0.020(0.032)	0.092(0.146)	0.037(0.046)
6	2.65(2.73)	0.11(0.10)	0.26(0.31)	0.05(0.05)	20.	0.133(0.184)	0.041(0.057)	0.190(0.289)	0.081(0.102)
7	4.33(4.80)	0.32(0.11)	0.45(0.32)	0.14(0.06)	120.	0.189(0.440)	0.026(0.096)	0.232(0.598)	0.151(0.283)
8	5.85(6.18)	0.26(0.17)	0.52(0.47)	0.13(0.08)	1000.	0.052(0.070)	0.003(0.030)	0.079(0.128)	0.034(0.035)
9	2.19(3.40)	0.86(0.17)	0.83(0.49)	0.42(0.09)	1000.	0.000(0.000)	0.003(0.000)	0.006(0.000)	0.000(0.000)
10	7.22(7.58)	0.18(0.09)	0.31(0.28)	0.09(0.05)	2200.	0.102(0.299)	0.020(0.076)	0.129(0.423)	0.069(0.189)
11	5.19(5.51)	0.34(0.24)	0.71(0.74)	0.16(0.13)	800.	0.049(0.263)	0.017(0.024)	0.083(0.106)	0.031(0.033)
12	7.06(7.24)	0.20(0.12)	0.40(0.40)	0.10(0.07)	1800.	0.175(0.242)	0.040(0.063)	0.244(0.335)	0.123(0.152)

β 0.58(0.38)   STD. ER. 0.15(0.05)   90% CONFIDENCE INTERVAL UPPER 0.94(0.43)   LOWER 0.48(0.28)

MODEL: Log N

TASK	TASK EFFECTS $\mu+v_k$	STD. ER.	TASK "STD. DEVS." σ_k	STD. ER.	WINDOW	TASK RISKS RISK	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
1	3.40(3.48)	0.25(0.21)	0.86(0.72)	0.16(0.12)	80.	0.128(0.120)	0.074(0.064)	0.256(0.227)	0.027(0.023)
2	3.71(3.78)	0.20(0.11)	0.67(0.43)	0.12(0.07)	600.	0.000(0.000)	0.043(0.034)	0.020(0.000)	0.000(0.004)
3	3.95(4.01)	0.18(0.15)	0.69(0.55)	0.12(0.10)	160.	0.051(0.042)	0.043(0.034)	0.148(0.112)	0.003(0.004)
4	5.94(5.93)	0.19(0.13)	0.60(0.47)	0.13(0.08)	100.	0.986(0.993)	0.026(0.008)	1.000(1.000)	0.918(0.973)
5	4.51(4.59)	0.18(0.13)	0.68(0.50)	0.12(0.09)	200.	0.123(0.109)	0.068(0.052)	0.262(0.203)	0.022(0.029)
6	2.62(2.65)	0.13(0.12)	0.52(0.52)	0.09(0.08)	20.	0.243(0.277)	0.092(0.083)	0.381(0.387)	0.089(0.130)
7	4.15(4.72)	0.11(0.10)	0.40(0.41)	0.08(0.07)	120.	0.312(0.448)	0.106(0.094)	0.473(0.610)	0.150(0.273)
8	6.17(6.11)	0.15(0.15)	0.52(0.52)	0.11(0.10)	1000.	0.067(0.089)	0.059(0.049)	0.177(0.174)	0.010(0.011)
9	3.44(3.43)	0.30(0.13)	0.77(0.45)	0.20(0.08)	1000.	0.000(0.000)	0.001(0.000)	0.000(0.000)	0.000(0.000)
10	7.50(7.61)	0.08(0.08)	0.26(0.20)	0.07(0.06)	2200.	0.277(0.404)	0.125(0.092)	0.481(0.547)	0.083(0.248)
11	5.41(5.42)	0.22(0.20)	0.93(0.83)	0.18(0.14)	800.	0.087(0.076)	0.054(0.041)	0.172(0.147)	0.013(0.018)
12	7.18(7.17)	0.14(0.15)	0.51(0.49)	0.10(0.09)	1800.	0.270(0.282)	0.097(0.098)	0.416(0.438)	0.109(0.120)

σ_ω 0.06(0.28)   STD. ER. 0.17(0.05)   90% CONFIDENCE INTERVAL UPPER 0.43(0.37)   LOWER 0.02(0.19)

Figures not in parentheses are estimates when no imputation is carried out for cases with missing values – methods described in the Appendices are employed; figures in parentheses are the corresponding estimates when values are imputed for missing cases.

TABLE 4

SYSTEM L – Missing only excluded

MODEL: Log EV

	TASK EFFECTS		TASK "STD. DEVS."		TASK RISKS			90% CONFIDENCE INTERVAL	
TASK	η_k	STD. ER.	ξ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	3.51(3.69)	0.26(0.20)	0.65(0.66)	0.13(0.12)	80.	0.102(0.134)	0.030(0.044)	0.151(0.224)	0.063(0.063)
2	3.80(4.24)	0.38(0.23)	0.90(0.80)	0.20(0.14)	600.	0.006(0.005)	0.002(0.003)	0.010(0.012)	0.004(0.002)
3	3.77(4.04)	0.25(0.24)	0.65(0.62)	0.14(0.10)	160.	0.033(0.045)	0.010(0.021)	0.052(0.086)	0.022(0.021)
4	5.32(5.96)	0.49(0.16)	0.72(0.44)	0.27(0.08)	100.	0.710(0.955)	0.084(0.045)	0.839(1.007)	0.592(0.882)
5	4.30(4.63)	0.25(0.18)	0.55(0.49)	0.11(0.09)	200.	0.049(0.080)	0.015(0.029)	0.079(0.139)	0.034(0.043)
6	2.64(2.72)	0.10(0.12)	0.29(0.33)	0.05(0.05)	20.	0.121(0.183)	0.031(0.065)	0.177(0.284)	0.075(0.095)
7	4.36(4.85)	0.26(0.11)	0.26(0.34)	0.13(0.05)	120.	0.184(0.481)	0.025(0.103)	0.231(0.657)	0.148(0.352)
8	5.90(6.13)	0.27(0.16)	0.50(0.46)	0.13(0.08)	1000.	0.035(0.046)	0.010(0.020)	0.055(0.089)	0.024(0.024)
9	3.14(4.75)	1.05(0.44)	1.87(1.24)	0.63(0.23)	1000.	0.034(0.040)	0.010(0.022)	0.053(0.085)	0.020(0.018)
10	7.23(7.58)	0.16(0.09)	0.35(0.31)	0.09(0.05)	2200.	0.103(0.103)	0.015(0.031)	0.129(0.131)	0.082(0.074)
11	5.26(5.57)	0.33(0.26)	0.81(0.79)	0.17(0.15)	800.	0.052(0.074)	0.016(0.023)	0.089(0.115)	0.034(0.041)
12	7.09(7.26)	0.16(0.14)	0.44(0.41)	0.10(0.07)	1800.	0.179(0.249)	0.042(0.073)	0.259(0.382)	0.121(0.136)

β

	STD. ER.		90% CONFIDENCE INTERVAL	
			UPPER	LOWER
0.42(0.33)	0.10(0.04)		0.49(0.36)	0.26(0.23)

MODEL: Log N

	TASK EFFECTS		TASK "STD. DEVS."		TASK RISKS			90% CONFIDENCE INTERVAL	
TASK	$\mu + v_k$	STD. ER.	σ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	3.56(3.60)	0.25(0.23)	0.91(0.81)	0.16(0.14)	80.	0.183(0.182)	0.085(0.082)	0.339(0.325)	0.051(0.049)
2	4.13(4.20)	0.23(0.18)	0.94(0.79)	0.15(0.12)	600.	0.008(0.004)	0.014(0.006)	0.039(0.017)	0.000(0.000)
3	3.94(4.00)	0.18(0.16)	0.69(0.57)	0.11(0.10)	160.	0.051(0.046)	0.049(0.036)	0.162(0.121)	0.003(0.006)
4	5.94(5.91)	0.19(0.13)	0.60(0.47)	0.12(0.09)	100.	0.987(0.991)	0.031(0.010)	1.000(1.000)	0.898(0.970)
5	4.50(4.59)	0.18(0.14)	0.68(0.53)	0.12(0.10)	200.	0.123(0.120)	0.069(0.055)	0.261(0.220)	0.022(0.037)
6	2.62(2.64)	0.13(0.14)	0.52(0.52)	0.09(0.09)	20.	0.239(0.275)	0.094(0.082)	0.383(0.387)	0.089(0.129)
7	4.59(4.78)	0.11(0.12)	0.40(0.44)	0.08(0.07)	120.	0.312(0.492)	0.103(0.096)	0.469(0.655)	0.152(0.315)
8	6.12(6.08)	0.14(0.12)	0.52(0.49)	0.11(0.10)	1000.	0.066(0.072)	0.062(0.043)	0.176(0.145)	0.003(0.006)
9	4.24(4.70)	0.55(0.31)	1.75(1.26)	0.33(0.20)	1000.	0.106(0.043)	0.071(0.033)	0.254(0.106)	0.011(0.005)
10	7.54(7.60)	0.09(0.08)	0.26(0.21)	0.07(0.07)	2200.	0.277(0.396)	0.125(0.092)	0.501(0.540)	0.082(0.240)
11	5.41(5.42)	0.22(0.21)	0.93(0.87)	0.18(0.15)	800.	0.087(0.084)	0.057(0.044)	0.186(0.159)	0.014(0.021)
12	7.18(7.19)	0.14(0.15)	0.51(0.48)	0.10(0.09)	1800.	0.270(0.295)	0.098(0.100)	0.414(0.462)	0.110(0.130)

σ_ω

	STD. ER.		90% CONFIDENCE INTERVAL	
			UPPER	LOWER
0.06(0.29)	0.22(0.06)		0.48(0.38)	0.02(0.18)

Figures not in parentheses are estimates when no imputation is carried out for cases with missing values – methods described in the Appendices are employed; figures in parentheses are the corresponding estimates when values are imputed for missing cases.

SYSTEM D — Missing only excluded

MODEL: Log EV

TASK	TASK EFFECTS η_k	STD. ER.	TASK "STD. DEVS." ξ_k	STD. ER.	WINDOW	TASK RISKS RISK	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
1	2.44(2.45)	0.17(0.17)	0.41(0.41)	0.09(0.10)	22.	0.027(0.029)	0.025(0.029)	0.085(0.091)	0.008(0.006)
2	5.75(5.75)	0.14(0.13)	0.34(0.34)	0.09(0.09)	450.	0.093(0.095)	0.047(0.048)	0.166(0.177)	0.037(0.035)
3	5.91(5.91)	0.09(0.10)	0.24(0.24)	0.06(0.06)	600.	0.007(0.008)	0.009(0.008)	0.027(0.019)	0.002(0.002)
4	5.42(5.70)	0.47(0.32)	1.00(0.81)	0.31(0.20)	1000.	0.037(0.036)	0.019(0.024)	0.075(0.088)	0.019(0.012)
5	6.96(7.03)	0.16(0.13)	0.38(0.34)	0.11(0.08)	2000.	0.022(0.021)	0.017(0.019)	0.059(0.066)	0.008(0.005)
6	5.77(5.78)	0.12(0.12)	0.26(0.26)	0.07(0.07)	600.	0.002(0.002)	0.004(0.005)	0.010(0.010)	0.000(0.000)
7	4.40(4.40)	0.23(0.22)	0.55(0.56)	0.14(0.14)	200.	0.025(0.025)	0.017(0.022)	0.065(0.066)	0.000(0.010)
8	5.76(5.77)	0.25(0.25)	0.66(0.67)	0.16(0.15)	1000.	0.019(0.021)	0.017(0.019)	0.048(0.057)	0.007(0.005)
9	6.13(6.14)	0.16(0.18)	0.44(0.44)	0.12(0.13)	1000.	0.018(0.019)	0.014(0.015)	0.043(0.051)	0.006(0.005)
10	6.93(6.93)	0.10(0.10)	0.26(0.26)	0.06(0.06)	1700.	0.009(0.009)	0.009(0.017)	0.025(0.027)	0.002(0.001)

β	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
0.17(0.17)	0.05(0.05)	0.27(0.28)	0.11(0.12)

MODEL: Log N

TASK	TASK EFFECTS $\mu + \nu_k$	STD. ER.	TASK "STD. DEVS." σ_k	STD. ER.	WINDOW	TASK RISKS RISK	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
1	2.37(2.37)	0.12(0.11)	0.43(0.39)	0.10(0.10)	22.	0.049(0.036)	0.043(0.036)	0.125(0.100)	0.000(0.001)
2	5.61(5.61)	0.15(0.17)	0.43(0.50)	0.09(0.11)	450.	0.127(0.161)	0.093(0.102)	0.325(0.352)	0.014(0.029)
3	5.87(5.87)	0.07(0.08)	0.22(0.24)	0.05(0.08)	600.	0.011(0.016)	0.040(0.029)	0.104(0.086)	0.000(0.000)
4	5.50(5.55)	0.32(0.26)	1.04(0.93)	0.24(0.20)	1000.	0.087(0.073)	0.064(0.056)	0.185(0.176)	0.002(0.004)
5	5.89(6.91)	0.14(0.14)	0.45(0.51)	0.10(0.11)	2000.	0.061(0.092)	0.049(0.057)	0.147(0.183)	0.001(0.007)
6	5.69(5.69)	0.07(0.09)	0.22(0.29)	0.06(0.09)	600.	0.001(0.010)	0.021(0.025)	0.034(0.076)	0.002(0.001)
7	4.27(4.27)	0.23(0.22)	0.73(0.71)	0.18(0.17)	200.	0.081(0.075)	0.062(0.058)	0.191(0.185)	0.002(0.001)
8	5.64(5.64)	0.25(0.23)	0.81(0.73)	0.21(0.19)	1000.	0.059(0.043)	0.052(0.043)	0.158(0.138)	0.001(0.000)
9	6.06(6.06)	0.15(0.12)	0.43(0.36)	0.09(0.09)	1000.	0.025(0.011)	0.048(0.028)	0.135(0.079)	0.001(0.000)
10	6.86(6.86)	0.11(0.09)	0.34(0.29)	0.08(0.09)	1700.	0.050(0.028)	0.047(0.036)	0.149(0.091)	0.003(0.001)

σ_ω	STD. ER.	90% CONFIDENCE INTERVAL UPPER	LOWER
0.06(0.07)	0.09(0.05)	0.21(0.20)	0.01(0.01)

Figures not in parentheses are estimates when no imputation is carried
out for cases with missing values - methods described in the Appendices
are employed; figures in parentheses are the corresponding estimates
when values are imputed for missing cases.

TABLE 6

SYSTEM D – Missing non-standard, excluded

MODEL: Log EV

	TASK EFFECTS		TASK "STD. DEVS."			TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	η_k	STD. ER.	ξ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	2.42(2.44)	0.18(0.17)	0.42(0.43)	0.10(0.10)	22.	0.040(0.043)	0.024(0.030)	0.087(0.117)	0.017(0.012)
2	5.73(5.73)	0.13(0.15)	0.33(0.34)	0.08(0.09)	450.	0.094(0.099)	0.050(0.053)	0.185(0.184)	0.042(0.038)
3	5.87(5.88)	0.09(0.09)	0.73(0.24)	0.06(0.06)	600.	0.007(0.008)	0.012(0.023)	0.021(0.028)	0.003(0.001)
4	5.55(5.99)	0.55(0.22)	0.78(0.56)	0.37(0.16)	1000.	0.028(0.031)	0.017(0.019)	0.056(0.074)	0.014(0.008)
5	6.92(6.99)	0.17(0.12)	0.37(0.35)	0.11(0.08)	2000.	0.023(0.024)	0.017(0.019)	0.056(0.066)	0.010(0.006)
6	5.74(5.75)	0.12(0.12)	0.26(0.26)	0.06(0.07)	600.	0.003(0.003)	0.016(0.022)	0.010(0.015)	0.001(0.015)
7	4.37(4.39)	0.20(0.23)	0.53(0.55)	0.13(0.14)	200.	0.027(0.030)	0.016(0.022)	0.057(0.076)	0.012(0.011)
8	4.78(5.29)	0.48(0.11)	0.68(0.29)	0.31(0.06)	1000.	0.000(0.000)	0.003(0.000)	0.005(0.000)	0.009(0.000)
9	6.08(6.10)	0.18(0.18)	0.18(0.44)	0.12(0.13)	1000.	0.018(0.019)	0.012(0.016)	0.043(0.055)	0.009(0.005)
10	6.91(6.90)	0.09(0.10)	0.24(0.26)	0.06(0.06)	1700.	0.009(0.009)	0.007(0.009)	0.023(0.026)	0.004(0.001)

β	STD. ER.		STD. ER.		90% CONFIDENCE INTERVAL
					UPPER LOWER
0.24(0.22)		0.06(0.06)			0.33(0.37) 0.13(0.17)

MODEL: Log N

	TASK EFFECTS		TASK "STD. DEVS."			TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	$\mu + \nu_k$	STD. ER.	σ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	2.37(2.37)	0.12(0.11)	0.42(0.37)	0.10(0.10)	22.	0.048(0.034)	0.042(0.034)	0.123(0.094)	0.000(0.000)
2	5.61(5.61)	0.15(0.17)	0.44(0.49)	0.09(0.11)	450.	0.130(0.163)	0.094(0.102)	0.320(0.356)	0.012(0.027)
3	5.87(5.87)	0.08(0.09)	0.22(0.25)	0.06(0.08)	600.	0.012(0.032)	0.035(0.037)	0.073(0.115)	0.000(0.001)
4	5.94(5.98)	0.25(0.15)	0.68(0.45)	0.20(0.10)	1000.	0.078(0.023)	0.073(0.028)	0.199(0.086)	0.000(0.001)
5	6.89(6.90)	0.14(0.10)	0.46(0.52)	0.10(0.11)	2000.	0.064(0.096)	0.050(0.058)	0.153(0.190)	0.002(0.008)
6	5.69(5.69)	0.07(0.10)	0.22(0.30)	0.06(0.09)	600.	0.001(0.016)	0.019(0.028)	0.020(0.031)	0.000(0.002)
7	4.27(4.27)	0.22(0.22)	0.73(0.72)	0.18(0.17)	200.	0.082(0.079)	0.061(0.058)	0.187(0.185)	0.002(0.002)
8	5.18(5.22)	0.27(0.14)	0.64(0.41)	0.19(0.12)	1000.	0.004(0.000)	0.014(0.001)	0.035(0.002)	0.000(0.000)
9	6.06(6.06)	0.15(0.12)	0.42(0.35)	0.09(0.08)	1000.	0.024(0.011)	0.046(0.026)	0.117(0.074)	0.000(0.000)
10	6.86(6.86)	0.11(0.10)	0.34(0.30)	0.07(0.09)	1700.	0.048(0.038)	0.046(0.041)	0.143(0.122)	0.003(0.002)

σ_ω	STD. ER.		STD. ER.		90% CONFIDENCE INTERVAL
					UPPER LOWER
0.08(0.13)		0.08(0.06)			0.18(0.24) 0.01(0.05)

Figures not in parentheses are estimates when no imputation is carried
out for cases with missing values – methods described in the Appendices
are employed; figures in parentheses are the corresponding estimates
when values are imputed for missing cases.

MODEL: Log EV

	TASK EFFECTS		TASK "STD. DEVS."			TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	η_k	STD. ER.	ξ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	3.61(3.74)	0.22(0.19)	0.70(0.67)	0.13(0.12)	80.	0.118(0.142)	0.035(0.043)	0.184(0.222)	0.074(0.079)
3	3.87(4.07)	0.22(0.16)	0.68(0.61)	0.13(0.11)	160.	0.034(0.041)	0.012(0.020)	0.054(0.073)	0.019(0.017)
5	4.39(4.64)	0.24(0.16)	0.58(0.49)	0.14(0.10)	200.	0.051(0.074)	0.020(0.038)	0.091(0.140)	0.028(0.027)
6	2.67(2.76)	0.10(0.09)	0.32(0.32)	0.07(0.06)	20.	0.154(0.191)	0.035(0.054)	0.189(0.281)	0.068(0.110)
7	4.40(4.80)	0.21(0.11)	0.46(0.32)	0.12(0.06)	120.	0.175(0.430)	0.034(0.102)	0.234(0.603)	0.127(0.273)
8	6.00(6.21)	0.21(0.14)	0.53(0.46)	0.12(0.08)	1000.	0.039(0.052)	0.015(0.028)	0.070(0.099)	0.024(0.021)
10	7.22(7.58)	0.22(0.11)	0.41(0.29)	0.13(0.06)	2200.	0.104(0.284)	0.023(0.084)	0.149(0.421)	0.074(0.155)
11	5.35(5.54)	0.27(0.22)	0.82(0.77)	0.16(0.16)	800.	0.045(0.056)	0.019(0.026)	0.084(0.104)	0.027(0.026)
12	7.14(7.27)	0.17(0.13)	0.46(0.41)	0.10(0.08)	1800.	0.192(0.243)	0.044(0.073)	0.272(0.375)	0.129(0.143)

$\hat{\mu}$ STD. ER.

0.30(0.28)   0.07(0.06)

90% CONFIDENCE INTERVAL
UPPER   LOWER
0.43(0.37)   0.20(0.20)

Model: Log N

	TASK EFFECTS		TASK "STD. DEVS."			TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	$\mu + \nu_k$	STD. ER.	σ_k	STD. ER.	WINDOW	RISK	STD. ER.	UPPER	LOWER
1	3.56(3.60)	0.23(0.22)	0.91(0.84)	0.16(0.15)	80.	0.183(0.189)	0.078(0.077)	0.313(0.308)	0.052(0.058)
3	3.94(4.01)	0.17(0.14)	0.69(0.56)	0.14(0.12)	160.	0.051(0.043)	0.042(0.034)	0.139(0.116)	0.002(0.004)
5	4.50(4.59)	0.17(0.13)	0.69(0.51)	0.11(0.09)	200.	0.124(0.111)	0.064(0.056)	0.233(0.212)	0.029(0.028)
6	2.63(2.65)	0.15(0.16)	0.52(0.56)	0.10(0.11)	20.	0.238(0.289)	0.102(0.097)	0.448(0.500)	0.086(0.154)
7	4.59(4.74)	0.12(0.11)	0.40(0.38)	0.08(0.07)	120.	0.310(0.460)	0.119(0.101)	0.492(0.598)	0.105(0.252)
8	6.12(6.12)	0.14(0.14)	0.52(0.53)	0.11(0.11)	1000.	0.065(0.094)	0.051(0.056)	0.160(0.210)	0.003(0.014)
10	7.46(7.56)	0.11(0.09)	0.39(0.25)	0.09(0.07)	2200.	0.276(0.358)	0.102(0.097)	0.455(0.510)	0.128(0.204)
11	5.41(5.42)	0.23(0.20)	0.94(0.80)	0.16(0.13)	800.	0.087(0.068)	0.048(0.040)	0.166(0.144)	0.012(0.010)
12	7.18(7.18)	0.13(0.13)	0.51(0.54)	0.09(0.08)	1800.	0.271(0.301)	0.095(0.083)	0.435(0.431)	0.109(0.175)

σ_ω STD. ER.

0.05(0.27)   0.14(0.07)

90% CONFIDENCE INTERVAL
UPPER   LOWER
0.35(0.36)   0.02(0.13)

Figures not in parentheses are estimates when no imputation is carried out for cases with missing values – methods described in the Appendices are employed; figures in parentheses are the corresponding estimates when values are imputed for missing cases.

TABLE 8

SYSTEM L - Missing non-standard, excluded
         - Tasks 2,4,9 excluded

MODEL: Log EV

	TASK EFFECTS		TASK "STD. DEVS."		WINDOW	TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	η_k	STD. ER.	ξ_k	STD. ER.		RISK	STD. ER.	UPPER	LOWER
1	3.31(3.62)	0.28(0.18)	0.74(0.61)	0.17(0.12)	80.	0.067(0.089)	0.017(0.029)	0.098(0.143)	0.043(0.046)
3	3.87(4.08)	0.20(0.17)	0.67(0.61)	0.13(0.11)	160.	0.035(0.041)	0.014(0.020)	0.059(0.081)	0.020(0.018)
5	4.39(4.64)	0.23(0.17)	0.58(0.50)	0.14(0.10)	200.	0.054(0.078)	0.019(0.040)	0.093(0.158)	0.032(0.030)
6	2.68(2.76)	0.11(0.09)	0.32(0.32)	0.07(0.06)	20.	0.140(0.196)	0.038(0.052)	0.214(0.277)	0.092(0.117)
7	4.40(4.79)	0.21(0.11)	0.45(0.31)	0.12(0.06)	120.	0.174(0.415)	0.034(0.105)	0.240(0.586)	0.126(0.247)
8	6.00(6.21)	0.21(0.13)	0.53(0.45)	0.13(0.08)	1000.	0.109(0.051)	0.023(0.028)	0.067(0.103)	0.023(0.022)
10	7.31(7.62)	0.19(0.10)	0.33(0.28)	0.11(0.06)	2200.	0.108(0.323)	0.024(0.086)	0.155(0.480)	0.078(0.190)
11	5.35(5.55)	0.27(0.22)	0.81(0.76)	0.19(0.16)	800.	0.045(0.057)	0.021(0.026)	0.077(0.110)	0.024(0.027)
12	7.14(7.28)	0.17(0.12)	0.45(0.40)	0.11(0.08)	1800.	0.187(0.246)	0.048(0.070)	0.278(0.357)	0.117(0.150)

β		STD. ER.		90% CONFIDENCE INTERVAL	
				UPPER	LOWER
0.31(0.28)		0.07(0.06)		0.44(0.38)	0.22(0.19)

Model: Log N

	TASK EFFECTS		TASK "STD. DEVS."		WINDOW	TASK RISKS		90% CONFIDENCE INTERVAL	
TASK	$\mu + \nu_k$	STD. ER.	σ_k	STD. ER.		RISK	STD. ER.	UPPER	LOWER
1	3.40(3.47)	0.23(0.21)	0.86(0.81)	0.16(0.14)	80.	0.129(0.144)	0.074(0.065)	0.245(0.245)	0.021(0.038)
3	3.95(4.01)	0.17(0.17)	0.69(0.58)	0.14(0.12)	160.	0.052(0.051)	0.055(0.035)	0.178(0.118)	0.002(0.005)
5	4.50(4.59)	0.17(0.13)	0.68(0.51)	0.11(0.08)	200.	0.124(0.110)	0.072(0.053)	0.276(0.194)	0.036(0.030)
6	2.63(2.65)	0.15(0.16)	0.52(0.56)	0.10(0.11)	20.	0.240(0.293)	0.102(0.098)	0.453(0.503)	0.095(0.150)
7	4.59(4.72)	0.12(0.14)	0.39(0.37)	0.08(0.06)	120.	0.309(0.445)	0.119(0.100)	0.494(0.593)	0.117(0.251)
8	6.12(6.13)	0.14(0.14)	0.52(0.51)	0.11(0.10)	1000.	0.066(0.092)	0.071(0.054)	0.251(0.199)	0.007(0.017)
10	7.54(7.63)	0.07(0.07)	0.26(0.10)	0.07(0.06)	2200.	0.278(0.418)	0.110(0.102)	0.475(0.588)	0.138(0.270)
11	5.41(5.42)	0.23(0.21)	0.93(0.82)	0.16(0.14)	800.	0.086(0.073)	0.057(0.041)	0.194(0.143)	0.012(0.009)
12	7.18(7.18)	0.13(0.13)	0.51(0.50)	0.09(0.07)	1800.	0.270(0.293)	0.097(0.081)	0.439(0.417)	0.130(0.169)

σ_ω		STD. ER.		90% CONFIDENCE INTERVAL	
				UPPER	LOWER
0.06(0.29)		0.31(0.06)		0.65(0.36)	0.03(0.18)

Figures not in parentheses are estimates when no imputation is carried
out for cases with missing values - methods described in the Appendices
are employed; figures in parentheses are the corresponding estimates
when values are imputed for missing cases.
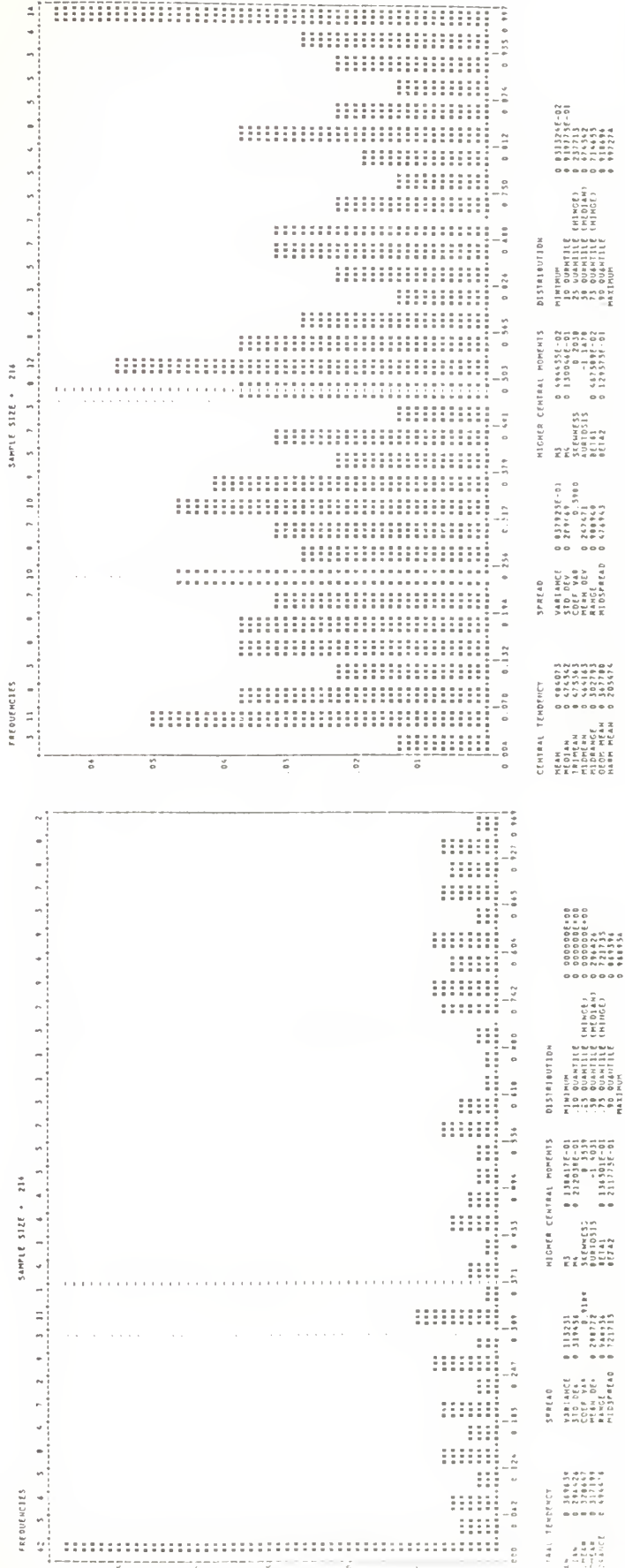
Figure 1. Uniform Residuals
System L, Model: Log N, no imputation
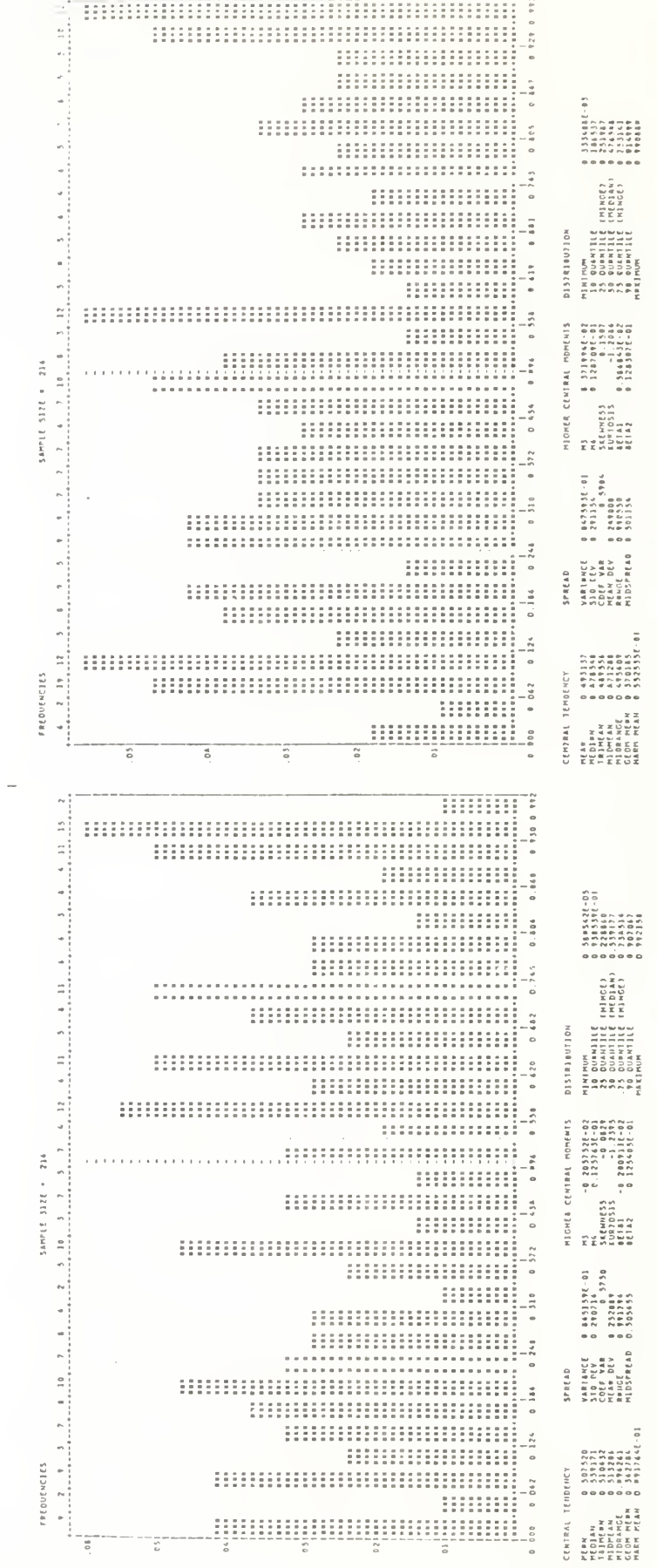
Figure 2. Uniform Residuals
System L, Model: Log N, imputation
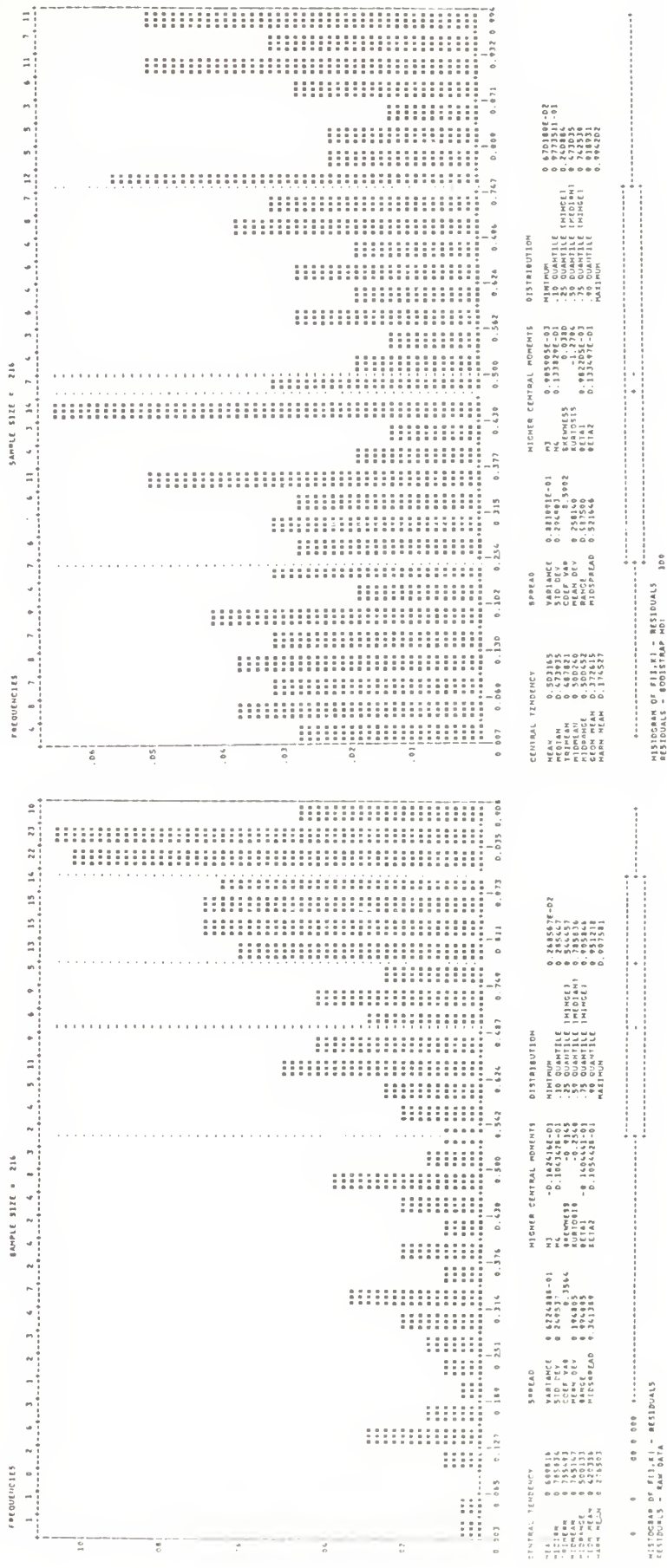
Figure 3. Uniform Residuals
System L, Model: Log EV, no imputation
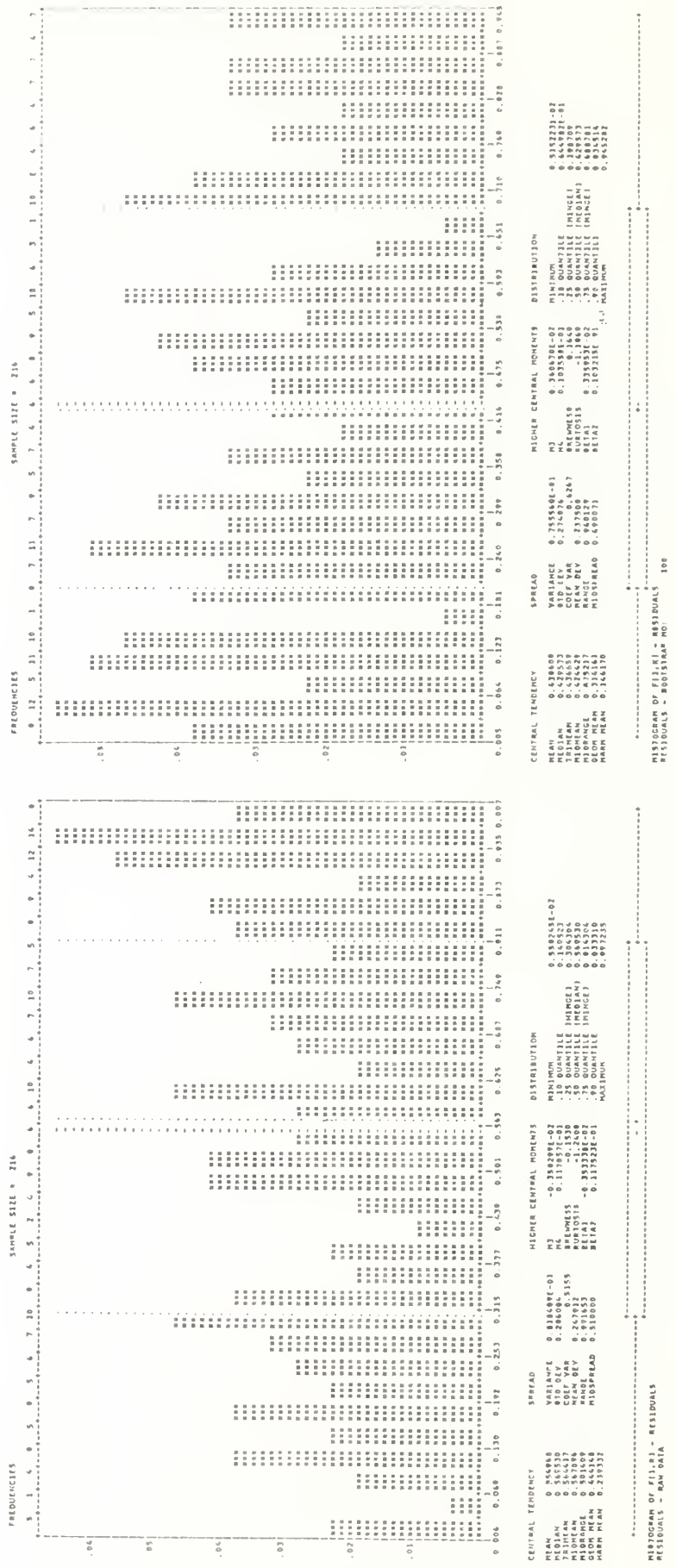
Figure 4. Uniform Residuals
System L, Model: Log EV, imputation
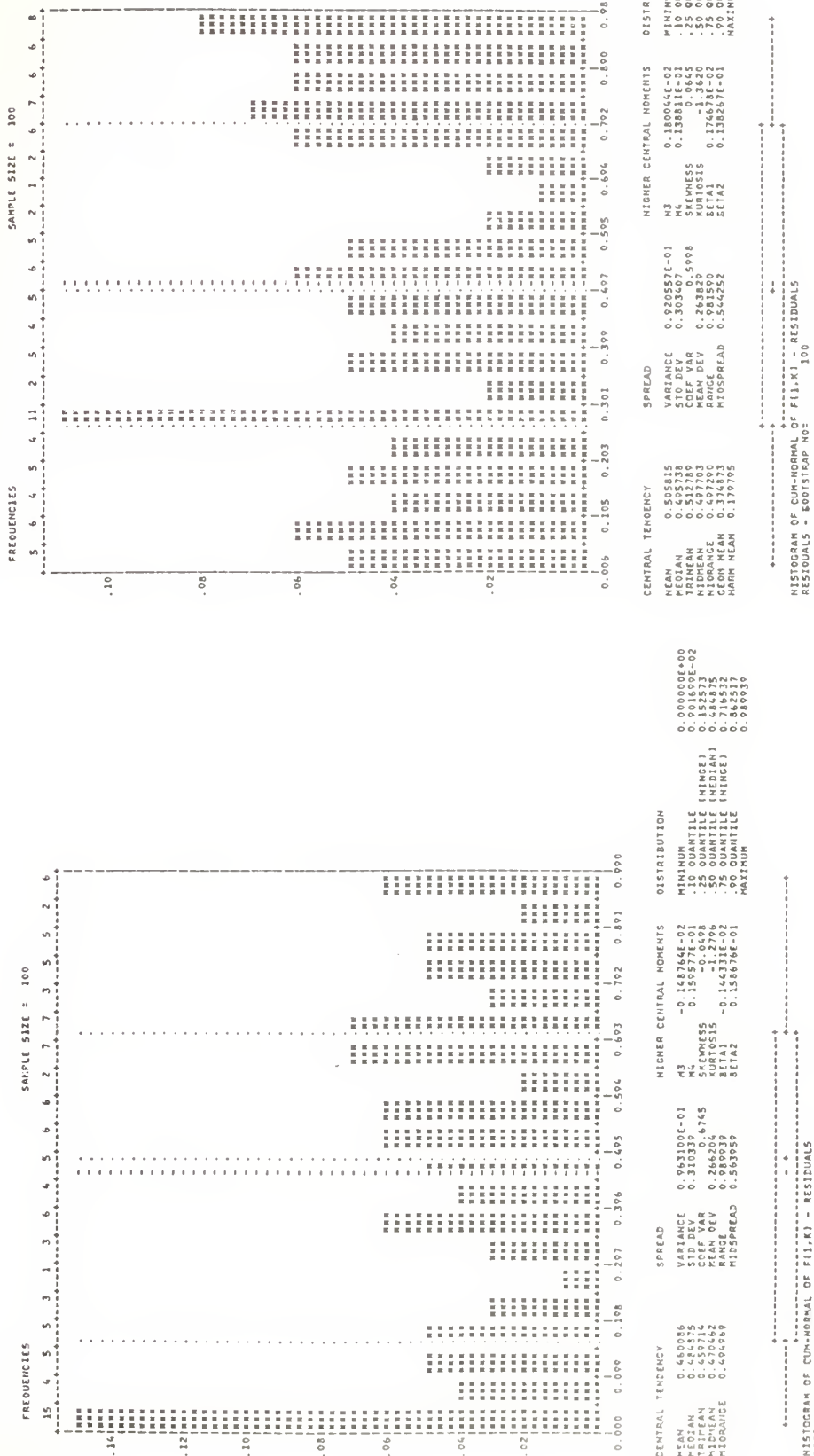
Figure 5. Uniform Residuals
System D, Model: Log N, no imputation
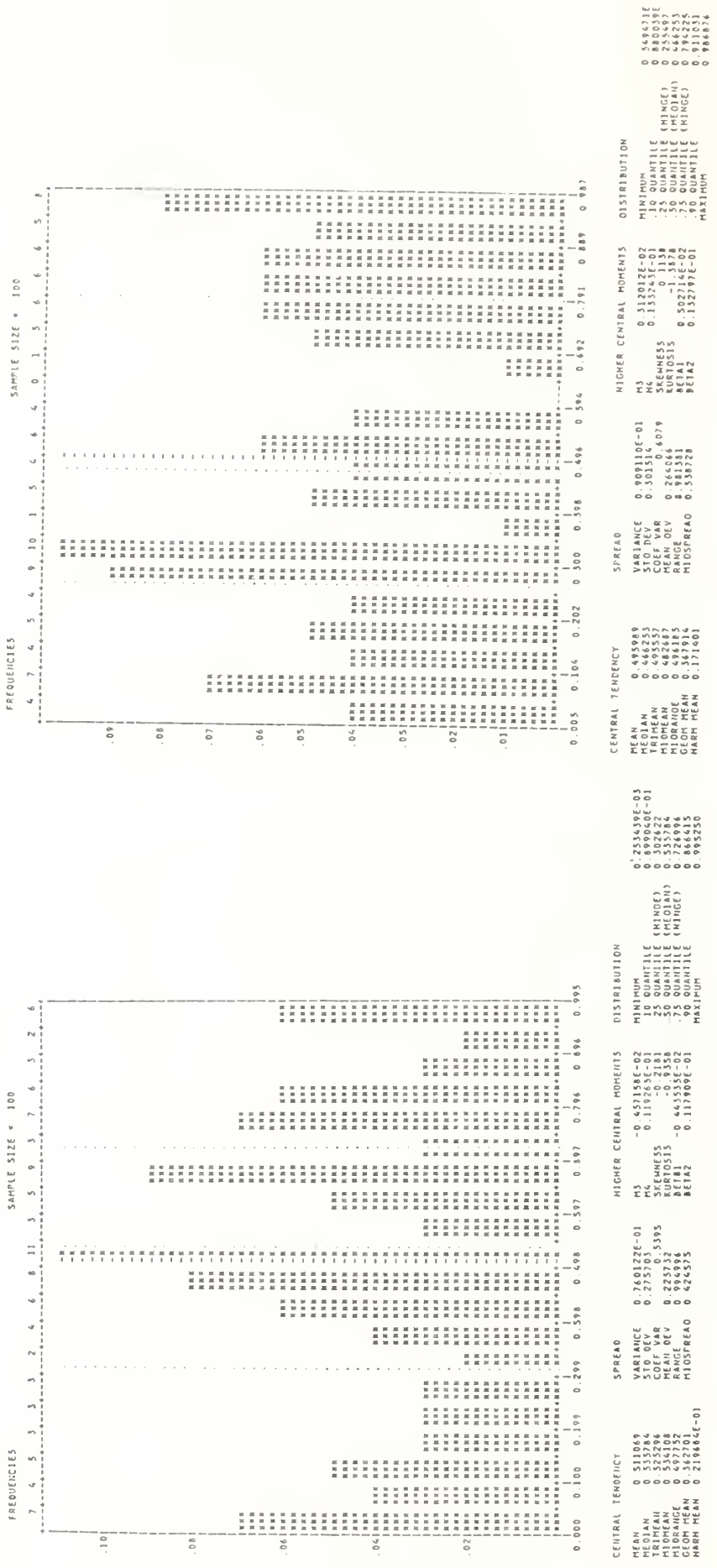
Figure 6. Uniform Residuals
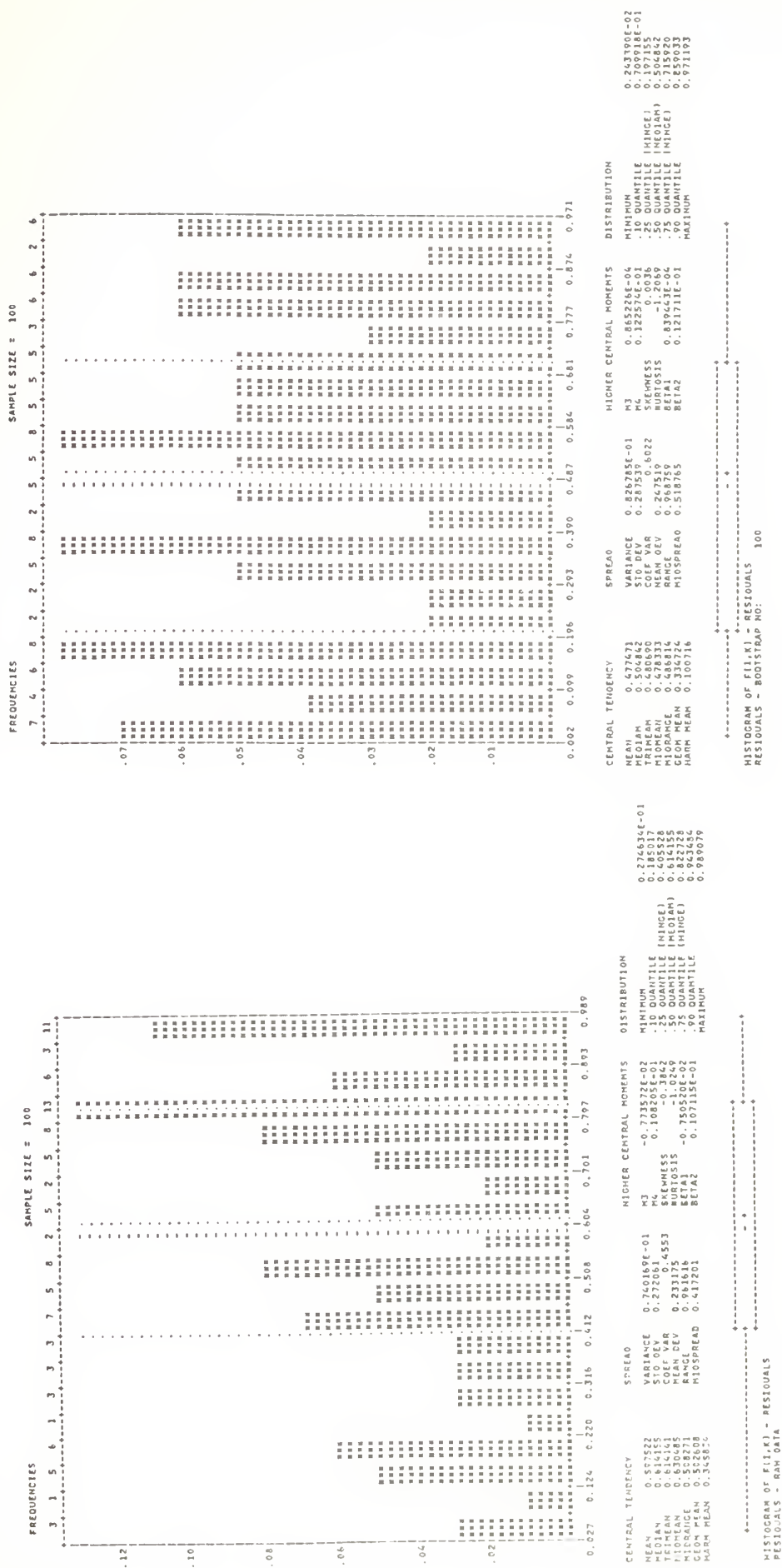
Figure 7. Uniform Residuals
System D, Model: Log EV, no imputation
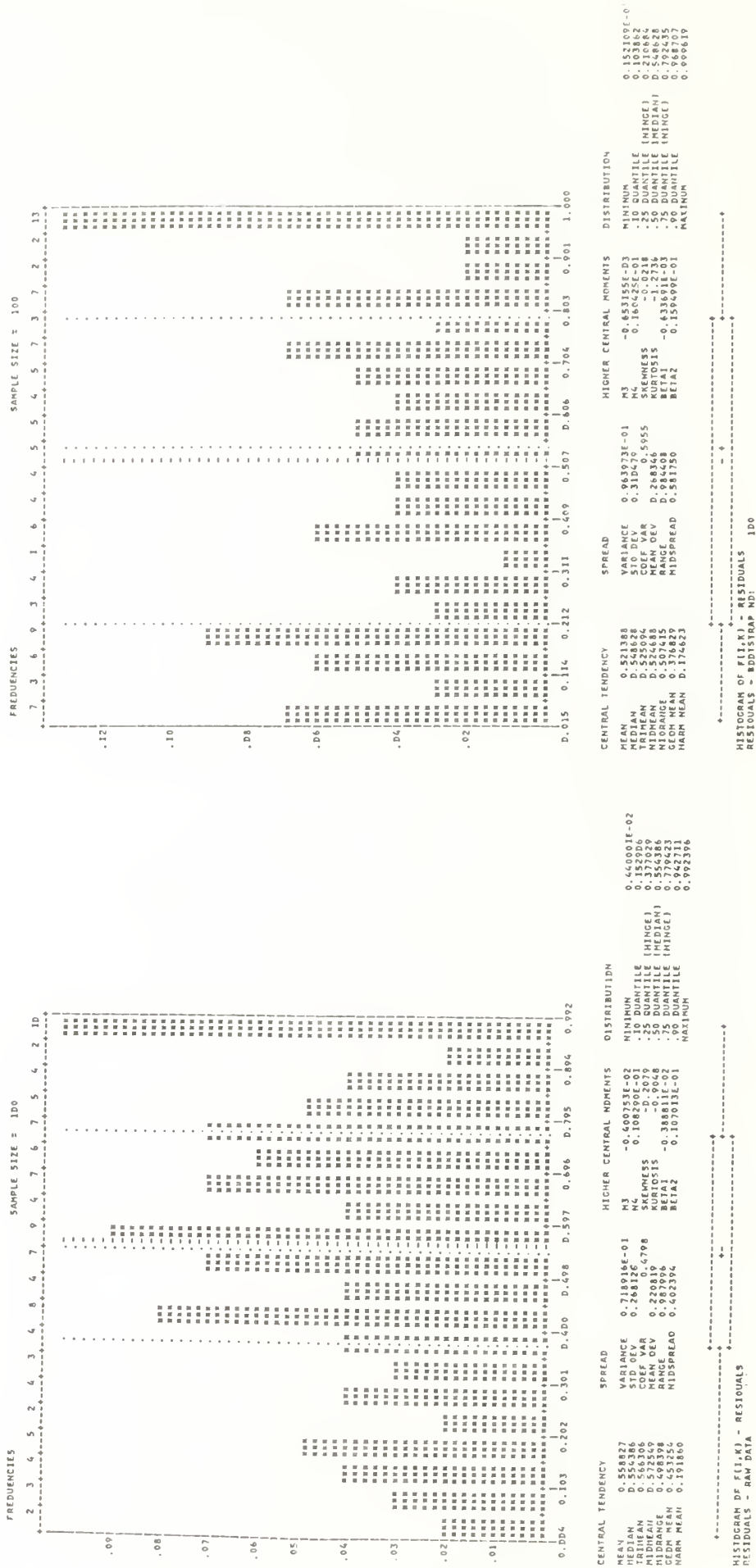
Figure 8. Uniform Residuals
System D, Model: Log EV, imputation

INITIAL DISTRIBUTION LIST

1. Library (Code 0142)...2
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Defense Technical Information Center...2
   Cameron Station
   Alexandria, VA 22314

3. Office of Research Administration (Code 012) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

4. Prof. Peter Purdue...1
   Code OR-Pd
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. Department of Operations Research (Code 55)............................1
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. Prof. Donald Gaver, Code OR-Gv..15
   Naval Postgraduate School
   Monterey, CA 93943-5000

7. Prof. Patricia Jacobs 1
   Code OR/Jc
   Naval Postgraduate School
   Monterey, CA 93943-5000

8. Center for Naval Analyses...1
   4401 Ford Avenue
   Alexandria, VA 22302-0268

9. Dr. David Brillinger..1
   Statistics Department
   University of California
   Berkeley, CA 94720

INITIAL DISTRIBUTION LIST

s

s

s

s